

Top 10 Words Per Item (index-based labels)

==== item_0 ===

commission: 5, said: 5, new: 4, proposal: 4, securities: 3, sec: 3, exchange: 2, agency: 2, understand: 2, impose: 2

==== item_1 ===

fed: 5, said: 3, federal: 2, reserve: 2, trust: 2, financial: 2, policy: 2, committee: 2, brown: 2, two: 1

==== item_2 ===

toll: 4, virginia: 3, said: 3, stuck: 2, traffic: 2, leesburg: 2, pike: 2, northern: 2, monday: 2, highway: 2

==== item_3 ===

tax: 7, code: 5, coalition: 4, overhaul: 4, corporate: 4, said: 3, boustany: 3, corporations: 2, nation: 2, microsoft: 2

==== item_4 ===

internet: 5, administration: 4, government: 3, said: 3, new: 2, even: 2, position: 2, issue: 2, would: 2, people: 2

==== item_5 ===

bill: 5, would: 4, debate: 3, iran: 3, authorize: 2, use: 2, force: 2, nuclear: 2, opponents: 2, war: 2

==== item_6 ===

said: 4, fed: 4, lot: 3, federal: 2, thursday: 2, banks: 2, would: 2, also: 2, securities: 2, financial: 2

==== item_7 ===

congress: 5, encryption: 5, export: 4, rules: 4, commission: 4, department: 3, products: 3, groups: 3, commerce: 2, new: 2

==== item_8 ===

said: 4, house: 4, encryption: 3, week: 3, letter: 3, government: 3, congress: 2, export: 2, policy: 2, protect: 2

==== item_9 ===

financial: 4, new: 4, banks: 4, fed: 4, federal: 3, bank: 3, regulators: 3, use: 3, technologies: 3, products: 3

==== item_10 ===

said: 3, central: 3, bank: 3, ruble: 2, fallen: 2, market: 2, dollar: 2, foreign: 2, currency: 2, reserves: 2

==== item_11 ===

bank: 5, central: 4, markets: 3, stock: 2, thursday: 2, said: 2, matijevic: 2, also: 2, seeking: 2, reserve: 2

==== item_12 ===

percent: 5, thursday: 3, gain: 2, index: 2, rose: 2, fell: 2, yields: 2, yield: 2, corporate: 2, fed: 2

==== item_13 ====
index: 6, percent: 4, euro: 4, stock: 2, exchange: 2, gain: 2, dollar: 2, czech: 2, republic: 2, year: 2

==== item_14 ====
percent: 5, prices: 4, november: 3, rose: 3, december: 3, czech: 2, consumer: 2, economists: 2, year: 2, bank: 2

==== item_15 ====
czech: 5, bill: 5, republic: 4, parties: 3, would: 3, european: 3, chamber: 3, register: 3, klaus: 2, parliament: 2

==== item_16 ====
said: 5, iran: 4, klaus: 3, agreement: 3, minister: 2, deal: 2, countries: 2, nuclear: 2, iranian: 2, zarif: 2

==== item_17 ====
likely: 3, billion: 3, years: 2, currency: 2, euro: 2, european: 2, euros: 2, government: 2, bonds: 2, assets: 2

==== item_18 ====
ecb: 4, said: 4, saunders: 3, stimulus: 2, effort: 2, euro: 2, inflation: 2, prague: 1, stock: 1, exchange: 1

==== item_19 ====
czech: 7, republic: 7, final: 4, entry: 4, song: 3, countries: 3, competition: 3, said: 3, may: 2, eurovision: 2

==== item_20 ====
commercial: 4, property: 4, market: 4, year: 4, according: 3, federation: 3, sector: 3, industry: 2, past: 2, small: 2

==== item_21 ====
bank: 9, ritchie: 6, former: 3, said: 3, complaint: 3, committee: 3, deutsche: 2, drop: 2, colleague: 2, conduct: 2

==== item_22 ====
foreign: 7, exchange: 7, cap: 6, bank: 5, per: 4, cent: 4, defend: 3, share: 3, trading: 3, european: 3

==== item_23 ====
fund: 6, superwoman: 3, horlick: 2, director: 2, said: 2, years: 2, cause: 2, always: 2, support: 2, abortion: 2

==== item_24 ====
fund: 5, said: 5, superwoman: 4, pension: 4, city: 4, horlick: 3, manager: 2, decision: 2, director: 2, hoplick: 2

==== item_25 ====

told: 4, years: 4, grenfell: 3, judge: 3, former: 2, stand: 2, trial: 2, manslaughter: 2, court: 2, action: 2

==== item_26 ===

bank: 6, focus: 4, grenfell: 3, people: 3, market: 3, tower: 2, new: 2, investors: 2, spokeswoman: 2, said: 2

==== item_27 ===

bank: 5, deutsche: 5, grenfell: 5, offices: 5, court: 5, uk: 4, decided: 3, sfo: 3, authorities: 3, breach: 3

==== item_28 ===

bank: 2, germany: 1, dresdner: 1, dutch: 1, abn: 1, amro: 1, holding: 1, nv: 1, also: 1, part: 1

==== item_29 ===

state: 6, investment: 5, banking: 4, firms: 3, giant: 2, merrill: 2, lynch: 2, expand: 2, bank: 2, citi: 2

==== item_30 ===

court: 6, lawyer: 5, liu: 4, top: 4, zhou: 4, lawsuit: 3, communist: 3, china: 3, legal: 3, said: 3

==== item_31 ===

china: 8, region: 5, hong: 4, kong: 4, special: 4, administrative: 4, said: 4, decision: 3, government: 3, territory: 2

==== item_32 ===

jiang: 4, china: 3, said: 3, new: 3, treaty: 3, united: 2, states: 2, nuclear: 2, weapons: 2, since: 2

==== item_33 ===

internet: 8, said: 5, government: 4, china: 3, open: 3, firms: 2, authority: 2, oversees: 2, chaplin: 2, chen: 2

==== item_34 ===

index: 5, china: 3, said: 3, markets: 3, shanghai: 3, composite: 3, record: 3, sunday: 2, chen: 2, surge: 2

==== item_35 ===

said: 5, us: 5, china: 4, materials: 4, chinese: 4, products: 4, move: 3, commerce: 3, department: 3, would: 3

==== item_36 ===

said: 5, interest: 3, rates: 3, inflation: 3, central: 3, bank: 3, fed: 3, warned: 2, higher: 2, meeting: 2

==== item_37 ===

foreign: 6, document: 4, draft: 3, firms: 3, markets: 2, investors: 2, companies: 2, subject: 2, rules: 2, neda: 2

==== item_38 ===

wang: 3, court: 3, chinese: 2, dissident: 2, trying: 2, online: 2, beijing: 2, said: 2, china: 2, tuesday: 2

==== item_39 ===

said: 8, court: 6, subversion: 4, wang: 4, charges: 4, sentenced: 3, prison: 3, spokesman: 3, jailed: 2, activists: 2

==== item_40 ===

news: 7, company: 6, corp: 5, stock: 3, results: 3, media: 2, percent: 2, sell: 2, stake: 2, global: 1

==== item_41 ===

loan: 7, growth: 6, bank: 5, home: 4, business: 4, cba: 3, per: 3, cent: 3, strong: 3, said: 3

==== item_42 ===

australia: 7, capital: 7, sale: 5, cba: 4, bank: 3, said: 3, properties: 3, asset: 3, laws: 3, home: 2

==== item_43 ===

economy: 4, percent: 4, quarter: 3, said: 2, pace: 2, first: 2, reuters: 2, poll: 2, survey: 2, economists: 2

==== item_44 ===

year: 7, billion: 6, bank: 4, quarter: 4, net: 3, income: 3, cba: 2, strong: 2, second: 2, positive: 2

==== item_45 ===

would: 7, new: 7, said: 6, paper: 6, images: 6, company: 4, newspaper: 3, include: 3, information: 3, statement: 3

==== item_46 ===

bank: 5, said: 4, policy: 4, murray: 3, low: 3, interest: 3, massive: 3, economy: 3, stimulus: 3, central: 2

==== item_47 ===

network: 5, seven: 4, murdoch: 4, australian: 3, deal: 3, company: 3, give: 3, billion: 2, would: 2, board: 2

==== item_48 ===

brambles: 4, said: 4, shipment: 3, government: 3, company: 3, pallet: 2, australian: 2, law: 2, export: 2, withdraw: 2

==== item_49 ===

government: 5, consortium: 4, cost: 4, project: 4, olympic: 3, stadium: 3, sydney: 3, state: 3, said: 3, building: 2

==== item_50 ===

said: 5, sell: 5, company: 4, looking: 3, businesses: 3, still: 3, buyer: 3, finance: 2, letter: 2, intent: 2

==== item_51 ===

deal: 5, said: 5, new: 4, mercury: 3, finance: 3, journal: 3, unions: 3, arizona: 2, person: 2, close: 2

==== item_52 ===

mercury: 4, stock: 4, investors: 3, bank: 3, company: 2, said: 2, million: 2, percent: 2, year: 2, market: 2

==== item_53 ===

said: 6, fbi: 4, statement: 4, san: 3, francisco: 3, individuals: 3, activities: 3, field: 2, office: 2, number: 2

==== item_54 ===

new: 6, said: 6, tax: 5, thornton: 4, paid: 3, companies: 3, would: 3, transactions: 2, mercury: 1, finance: 1

==== item_55 ===

percent: 6, nasdaq: 4, change: 3, dow: 3, rose: 3, points: 3, strong: 2, market: 2, stock: 2, jones: 2

==== item_56 ===

education: 6, said: 4, sherrod: 3, conference: 3, former: 2, people: 2, together: 2, message: 2, going: 2, debate: 2

==== item_57 ===

said: 5, employees: 5, mercury: 4, overtime: 3, pay: 3, commission: 3, company: 3, would: 3, finance: 2, thursday: 2

==== item_58 ===

bank: 7, company: 5, said: 3, ford: 3, manufacture: 2, dispute: 2, federal: 2, spokesman: 2, mercury: 1, finance: 1

==== item_59 ===

company: 6, said: 5, bank: 5, corp: 2, major: 2, stores: 2, reliance: 1, acceptance: 1, friday: 1, borrowed: 1

==== item_60 ===

said: 5, companies: 4, merger: 3, market: 3, billion: 3, announcement: 2, combined: 2, value: 2, offer: 2, number: 2

==== item_61 ===

police: 6, bissonnette: 5, said: 4, thursday: 3, two: 3, alexandre: 3, father: 3, expected: 2, charges: 2, disappearance: 2

==== item_62 ===

energy: 5, power: 4, sector: 3, customers: 3, said: 2, affected: 2, canadians: 1, experiencing: 1, winter: 1, dismay: 1

==== item_63 ===

new: 6, city: 5, law: 5, businesses: 5, also: 3, toronto: 2, requires: 2, said: 2, business: 2, largely: 2

==== item_64 ===

city: 6, ford: 4, toronto: 3, population: 3, growing: 2, problem: 2, said: 2, hennigan: 2, going: 2, need: 2

==== item_65 ===

mining: 4, indonesia: 4, said: 3, subianto: 3, foreign: 3, breakthrough: 2, efforts: 2, get: 2, companies: 2, invest: 2

==== item_66 ===

company: 6, meeting: 4, annual: 3, website: 3, year: 3, april: 2, new: 2, york: 2, report: 2, board: 2

==== item_67 ===

police: 8, man: 8, said: 5, vancouver: 4, officer: 4, provincial: 3, arrested: 3, british: 2, columbia: 2, head: 2

==== item_68 ===

city: 4, going: 3, strippers: 2, canada: 2, country: 2, public: 2, vancouver: 2, change: 2, said: 2, councillor: 2

==== item_69 ===

park: 4, service: 4, de: 3, guzman: 3, oil: 3, discovery: 3, national: 3, team: 3, geologist: 2, largest: 2

==== item_70 ===

said: 6, rule: 5, would: 5, safety: 4, cars: 4, used: 3, vehicles: 3, new: 3, agency: 3, motor: 2

==== item_71 ===

union: 6, auto: 4, workers: 3, chrysler: 3, said: 3, alliance: 3, canadian: 2, likely: 2, boost: 2, power: 2

==== item_72 ===

said: 5, million: 5, company: 3, pay: 3, gm: 3, settle: 2, ignition: 2, department: 2, general: 1, motors: 1

==== item_73 ===

shooting: 5, police: 5, said: 4, dealership: 3, louis: 3, greene: 3, mustang: 2, charged: 2, officer: 2, belmar: 2

==== item_74 ===

quaker: 7, breakfast: 4, lunch: 3, deal: 3, said: 3, customers: 3, oats: 2, billion: 2, market: 2, bring: 2

==== item_75 ===

kellogg: 7, breakfast: 4, perdue: 4, billion: 3, deal: 3, cereal: 3, people: 3, broaden: 2, offerings: 2, foods: 2

==== item_76 ===

workers: 8, deal: 3, said: 3, union: 3, ratified: 2, new: 2, american: 2, auto: 2, covers: 2, president: 2

==== item_77 ===

customers: 4, retailers: 3, selling: 3, study: 3, stores: 3, sales: 2, discount: 2, prices: 2, discounts: 2, retail: 2

==== item_78 ===

retail: 4, said: 4, sears: 3, morgan: 3, sales: 2, years: 2, share: 2, overall: 2, greenfield: 2, aggressive: 2

==== item_79 ===

chrysler: 6, billion: 4, said: 3, worth: 2, agreement: 2, administration: 2, fund: 2, government: 2, cash: 2, deal: 2

==== item_80 ===

deal: 5, would: 5, said: 4, company: 4, create: 4, aerospace: 3, largest: 3, lucasvarity: 2, defense: 2, parts: 2

==== item_81 ===

sales: 8, landline: 3, mobile: 2, televisions: 2, rise: 2, expected: 2, consumer: 2, said: 2, data: 2, association: 2

==== item_82 ===

fcc: 5, company: 4, blackwater: 3, sales: 3, products: 3, ragan: 3, communications: 2, block: 2, lawyer: 2, us: 2

==== item_83 ===

said: 5, solar: 5, racal: 4, modules: 4, would: 3, cell: 3, savings: 2, million: 2, plant: 2, bought: 2

==== item_84 ===

yen: 6, hanson: 5, said: 5, company: 4, billion: 3, percent: 3, rose: 3, reported: 2, million: 2, profit: 2

==== item_85 ===

billion: 8, revenue: 4, said: 4, rise: 3, pounds: 3, percent: 3, year: 3, rose: 3, hanson: 2, profit: 2

==== item_86 ===

said: 5, united: 4, states: 4, obama: 4, president: 3, regime: 3, congress: 3, britain: 2, bombing: 2, use: 2

==== item_87 ===

said: 6, government: 5, would: 5, icc: 4, british: 3, refer: 3, criminal: 3, ask: 3, investigate: 3, friday: 2

==== item_88 ===

government: 3, marijuana: 3, drug: 3, issue: 2, legalization: 2, trade: 2, canada: 2, minister: 2, policy: 2, going: 2

==== item_89 ===

british: 7, commission: 5, said: 5, european: 4, referendum: 4, government: 3, would: 3, country: 3, people: 3, eu: 3

==== item_90 ===

year: 7, billion: 7, full: 5, vlsi: 4, percent: 4, expects: 3, quarter: 3, revenue: 3, said: 3, fourth: 2

==== item_91 ===

percent: 6, company: 6, said: 4, billion: 4, stock: 3, vlsi: 3, friday: 2, chief: 2, executive: 2, white: 2

==== item_92 ===

rahami: 5, judge: 4, ordered: 4, ap: 3, held: 3, story: 2, longer: 2, wednesday: 2, new: 2, york: 2

==== item_93 ===

market: 4, fell: 3, points: 3, value: 2, day: 2, analysts: 2, predicting: 2, stock: 2, could: 2, would: 2

==== item_94 ===

said: 6, access: 4, computer: 3, gained: 3, email: 3, intrepid: 3, network: 2, hackers: 2, customer: 2, phishing: 2

==== item_95 ===

computer: 4, associates: 4, company: 3, software: 2, decision: 2, allow: 2, sell: 2, ammunition: 2, disappointed: 2, miller: 2

==== item_96 ===

rules: 5, new: 4, said: 3, data: 3, online: 3, companies: 3, proposal: 3, people: 3, corp: 2, internet: 2

==== item_97 ===

billion: 7, said: 5, net: 4, company: 3, loss: 3, million: 3, period: 3, shares: 2, trading: 2, quarterly: 2

==== item_98 ===

company: 4, said: 4, revenue: 4, percent: 4, billion: 4, office: 3, equipment: 3, rose: 3, year: 3, strong: 2

==== item_99 ===

monday: 3, going: 3, market: 3, selling: 2, pressure: 2, days: 2, think: 2, said: 2, mervis: 2, networking: 1

Stylistic Analysis

Total unique words: 3131

Total word count: 19066

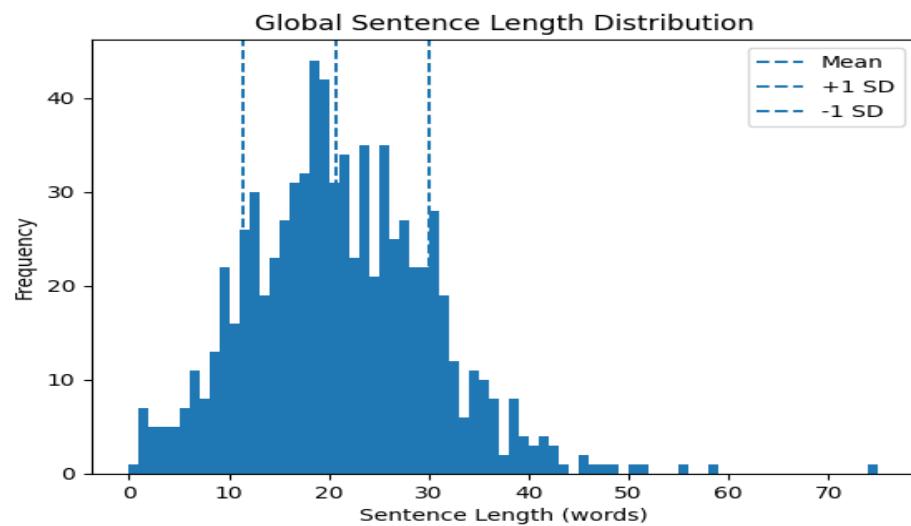
Type Token Ratio (TTR): 16.42%

Total stop-words count: 7067

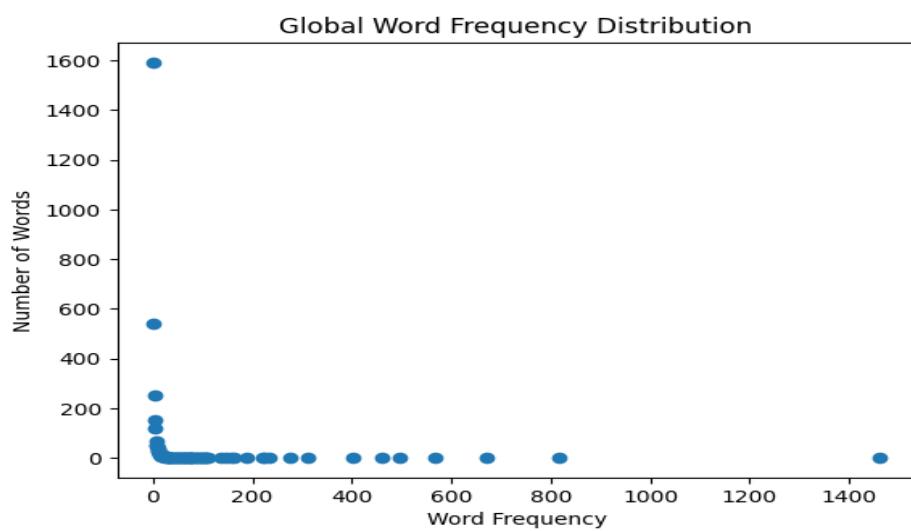
Most common Parts of Speech n-grams:

1. ('DT', 'NN'), 1305
2. ('NN', 'IN'), 962
3. ('IN', 'DT'), 896
4. ('JJ', 'NN'), 855
5. ('NN', 'NN'), 725
6. ('DT', 'JJ'), 588
7. ('NN', ','), 520
8. ('IN', 'NN'), 394
9. ('TO', 'VB'), 384
10. ('NN', 'VBD'), 380
11. ('NN', '.'), 360
12. ('.', 'DT'), 291
13. ('JJ', 'NNS'), 249
14. ('NNS', 'IN'), 239
15. ('IN', 'JJ'), 200

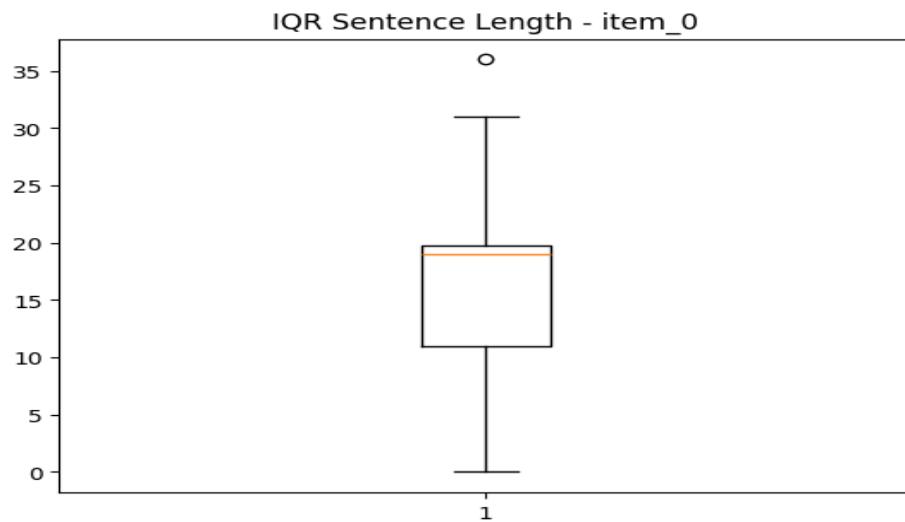
global_sentence_length_distribution.png



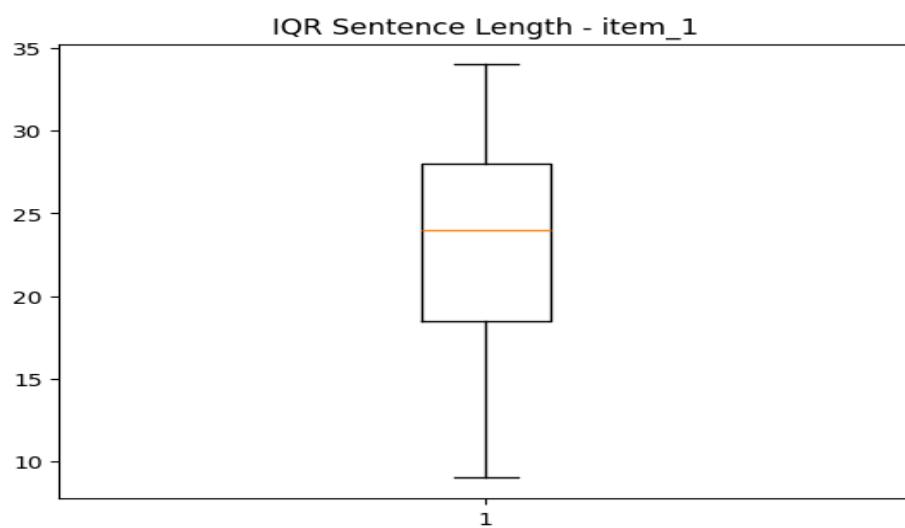
global_word_freq_distribution.png



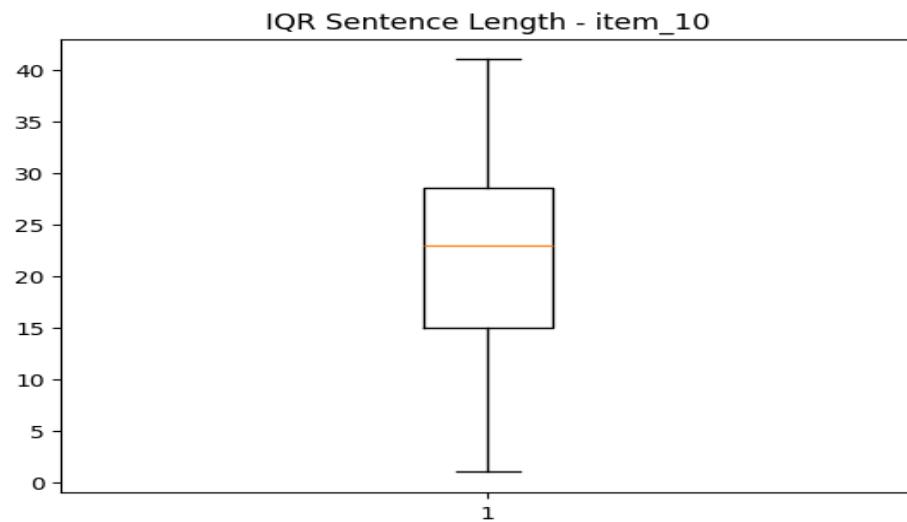
sentence_IQR_item_0.png



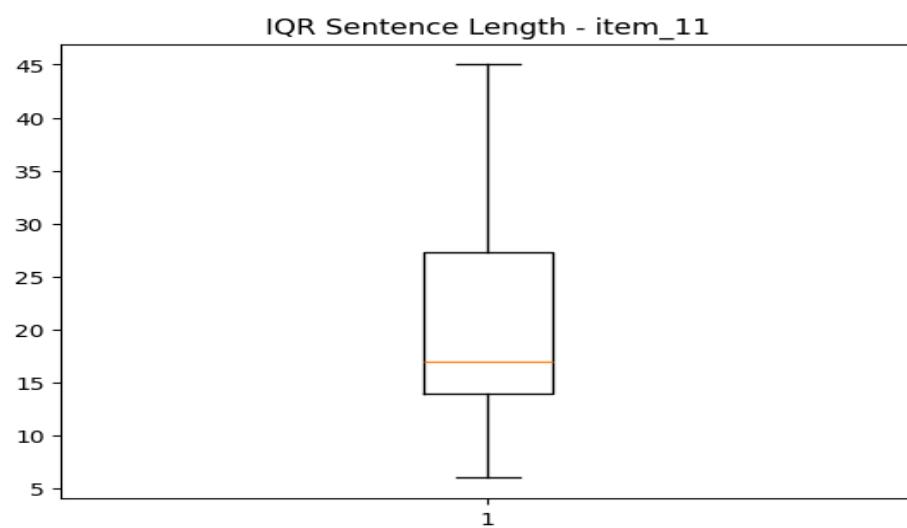
sentence_IQR_item_1.png



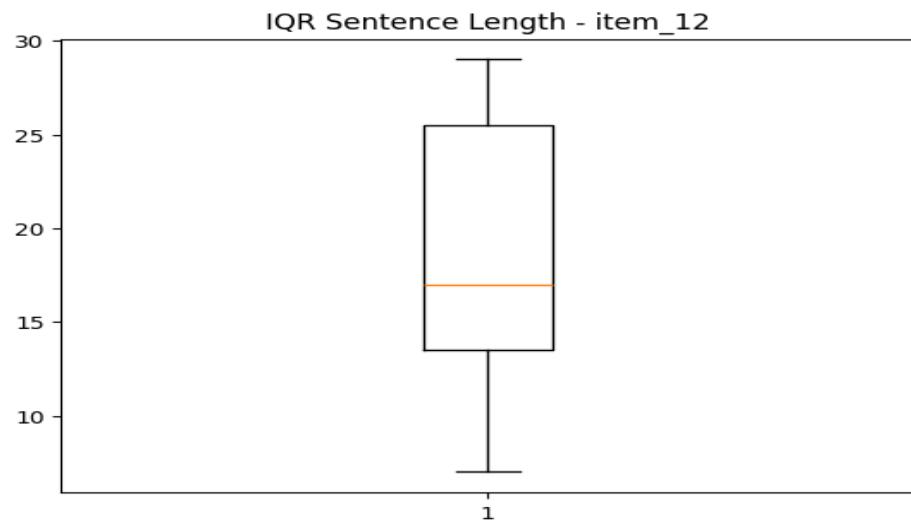
sentence_IQR_item_10.png



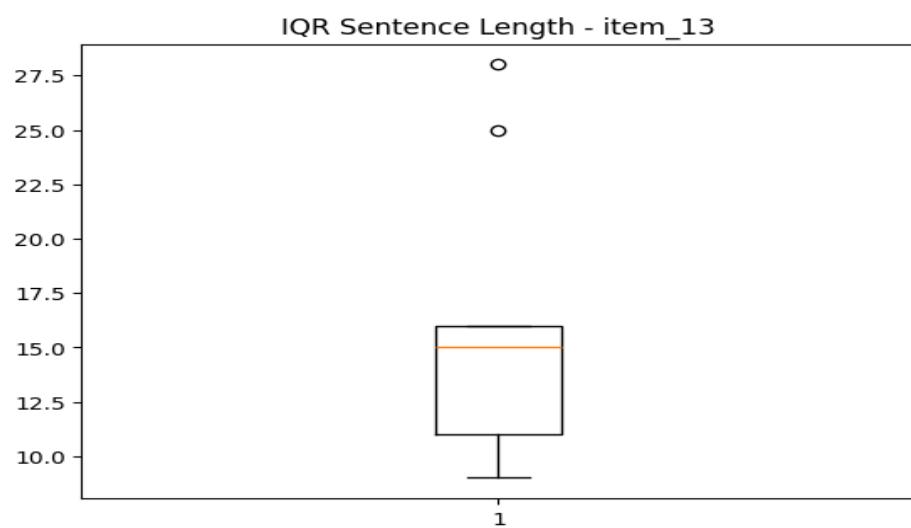
sentence_IQR_item_11.png



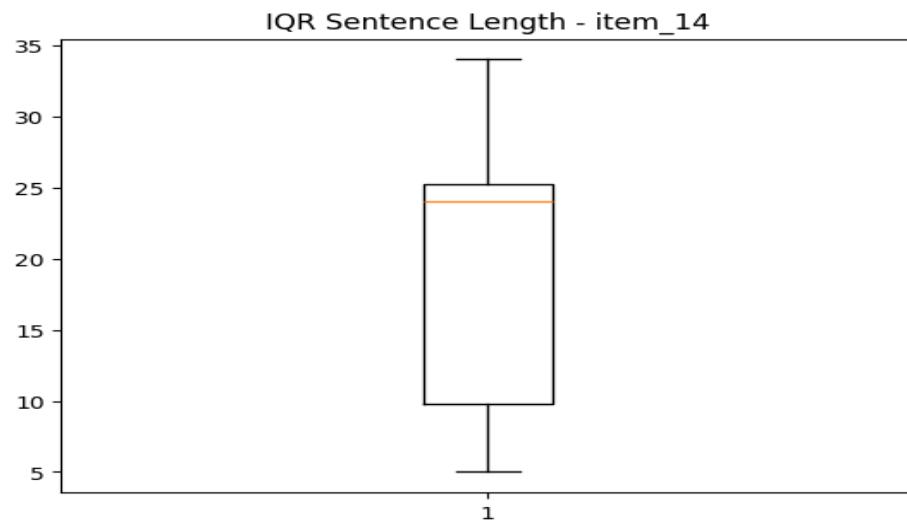
sentence_IQR_item_12.png



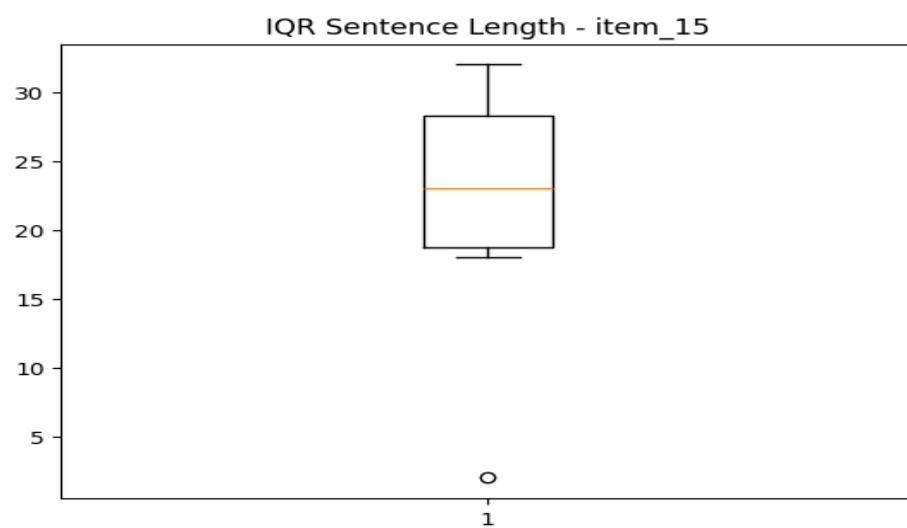
sentence_IQR_item_13.png



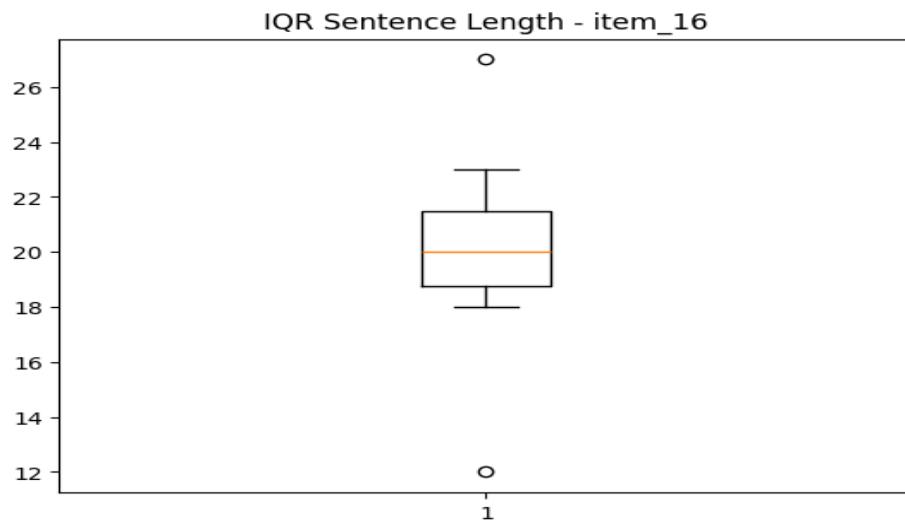
sentence_IQR_item_14.png



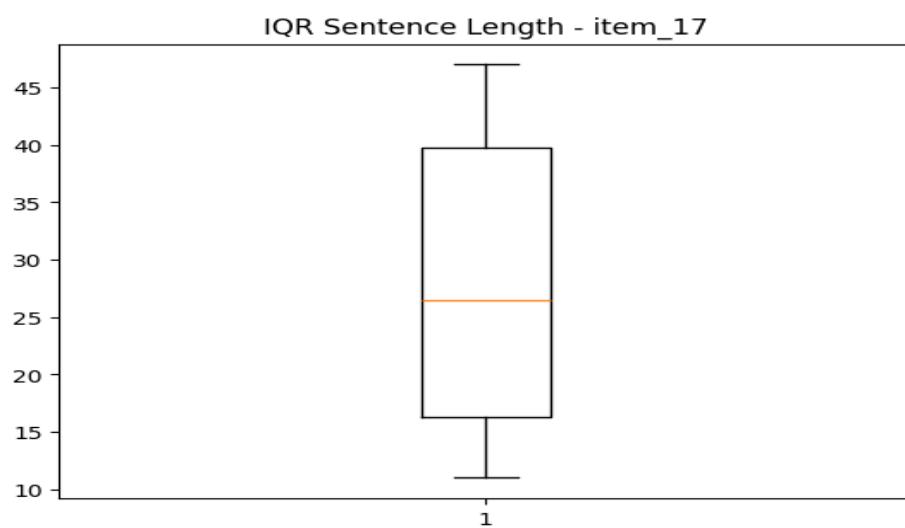
sentence_IQR_item_15.png



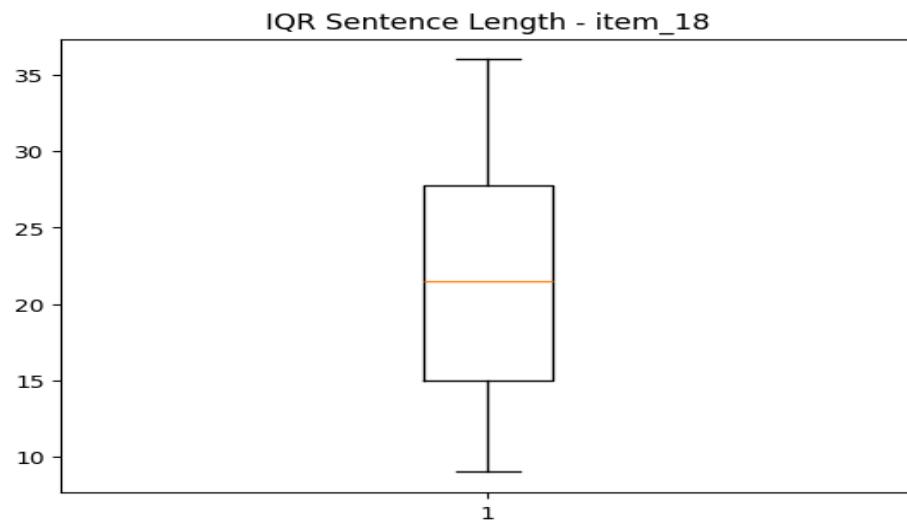
sentence_IQR_item_16.png



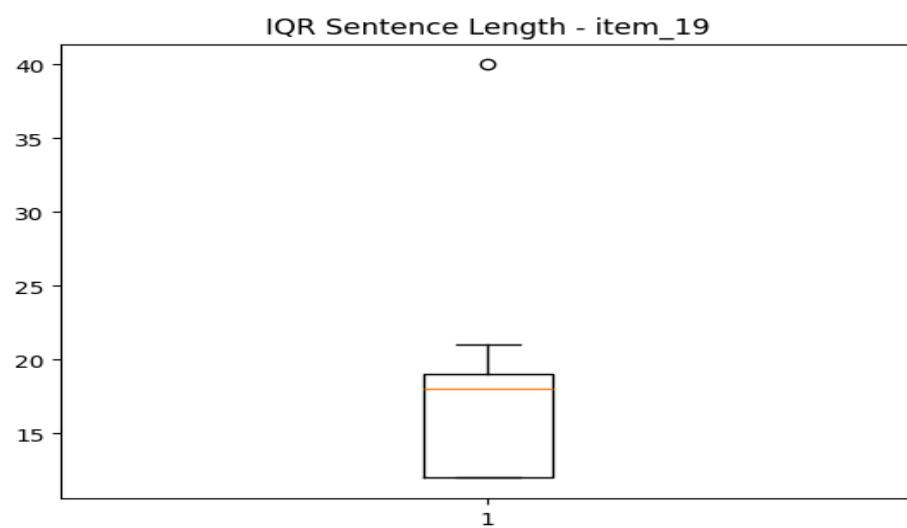
sentence_IQR_item_17.png



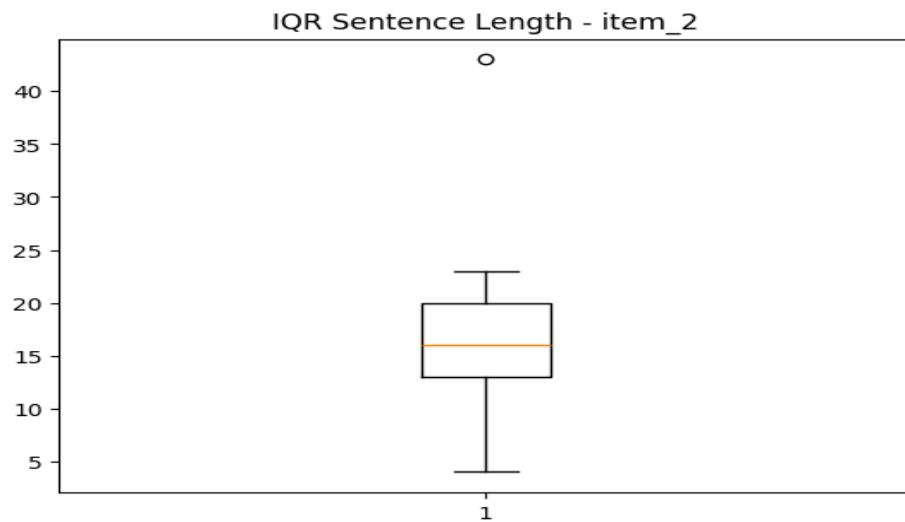
sentence_IQR_item_18.png



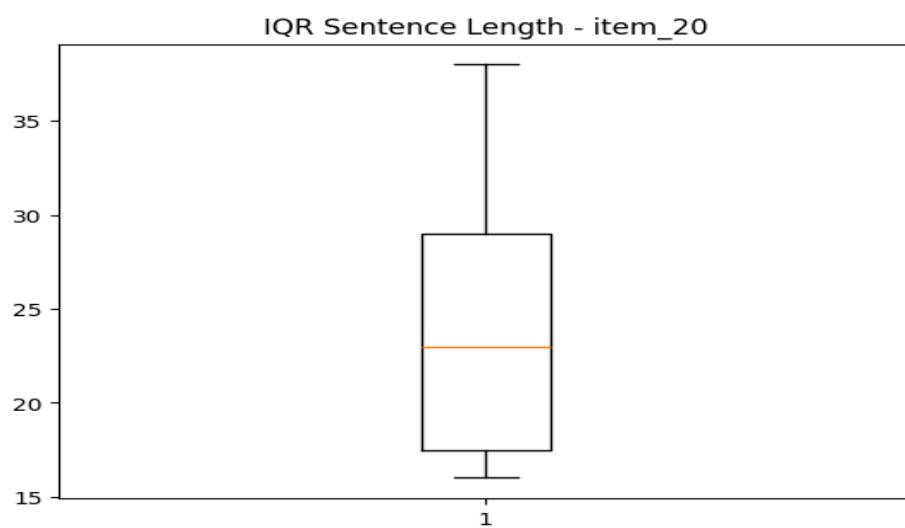
sentence_IQR_item_19.png



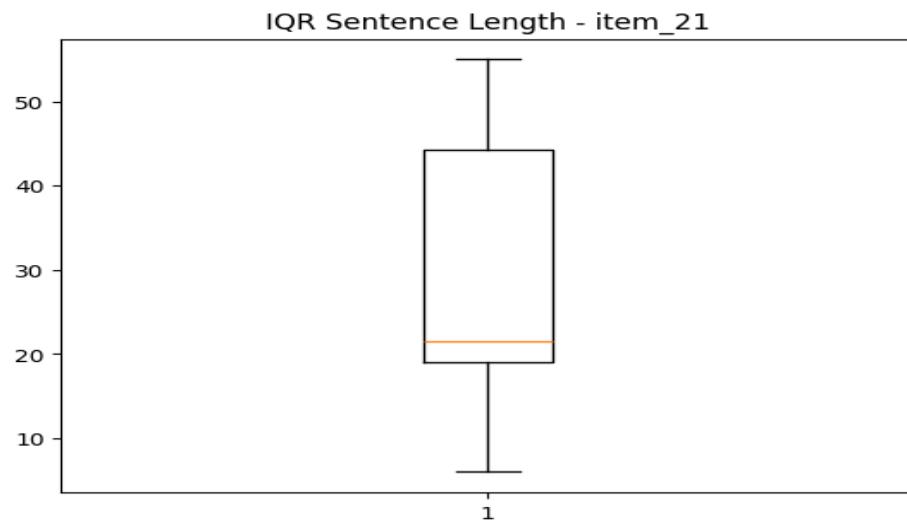
sentence_IQR_item_2.png



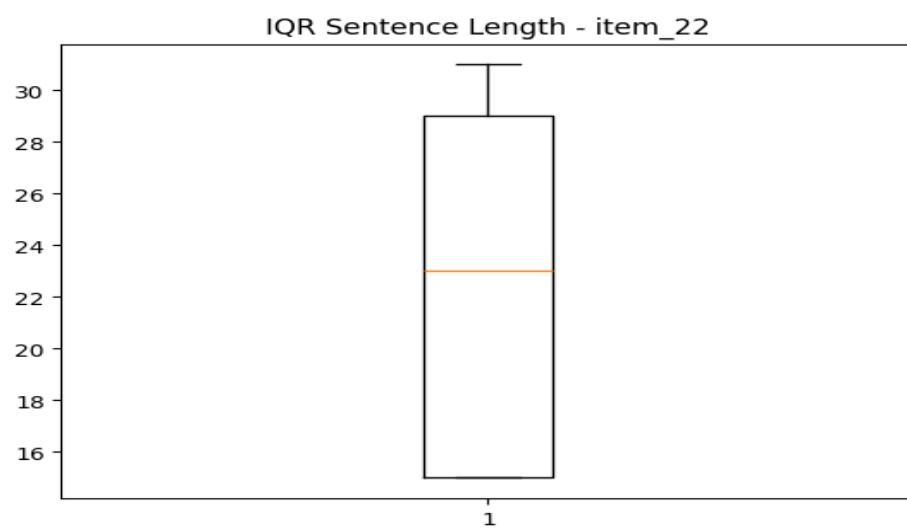
sentence_IQR_item_20.png



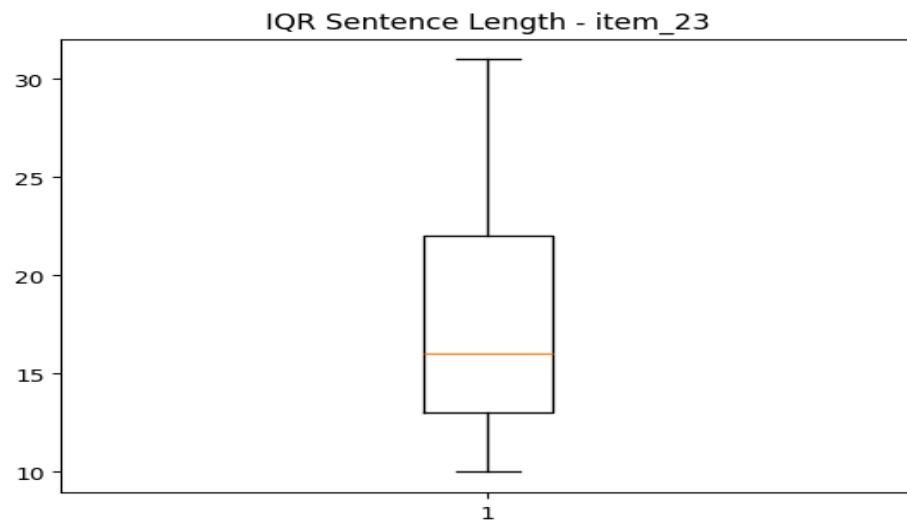
sentence_IQR_item_21.png



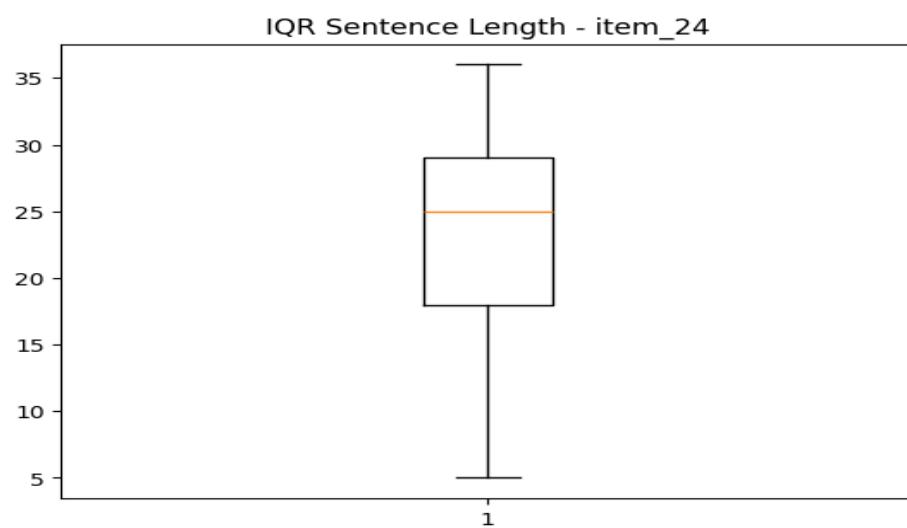
sentence_IQR_item_22.png



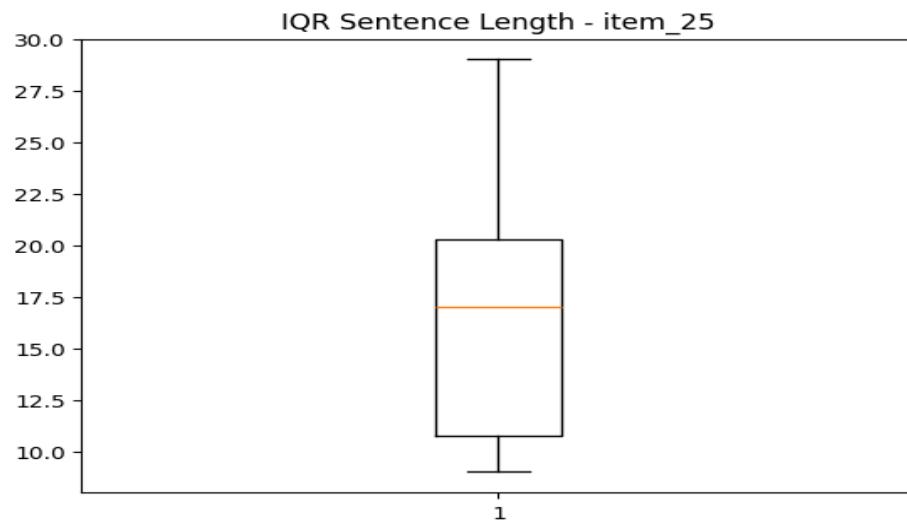
sentence_IQR_item_23.png



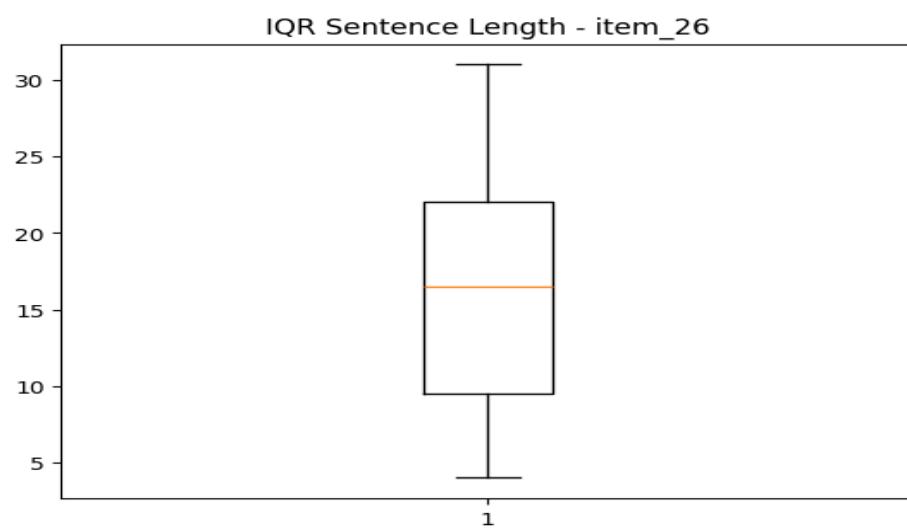
sentence_IQR_item_24.png



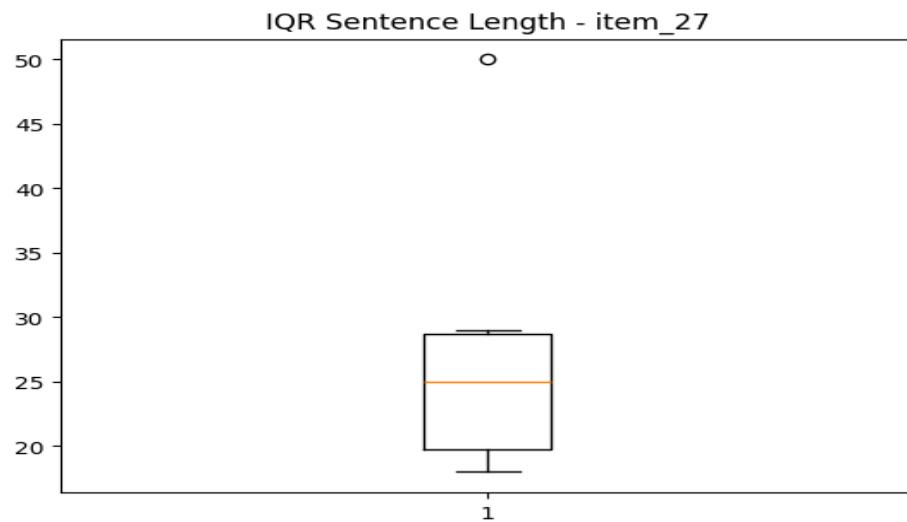
sentence_IQR_item_25.png



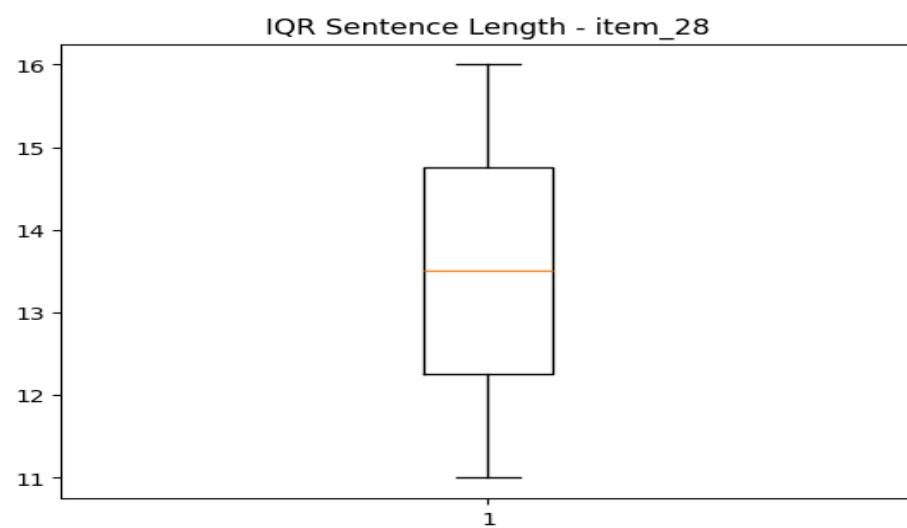
sentence_IQR_item_26.png



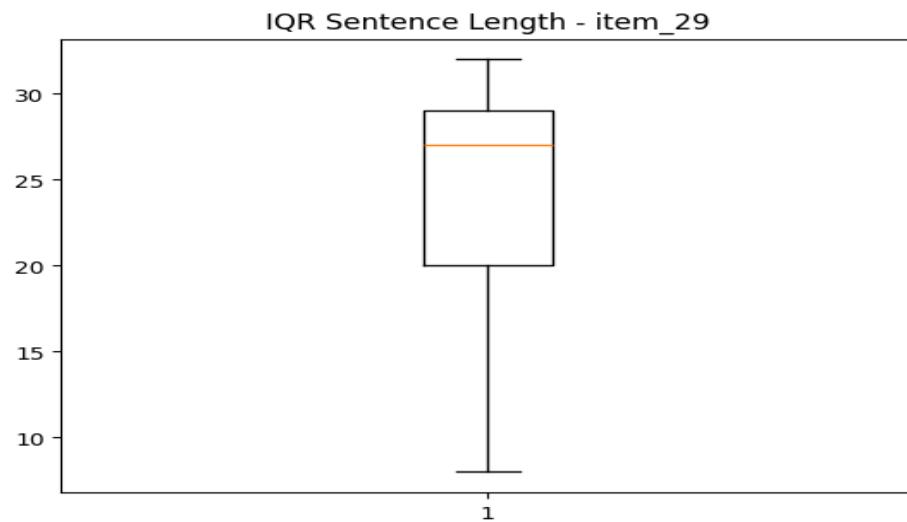
sentence_IQR_item_27.png



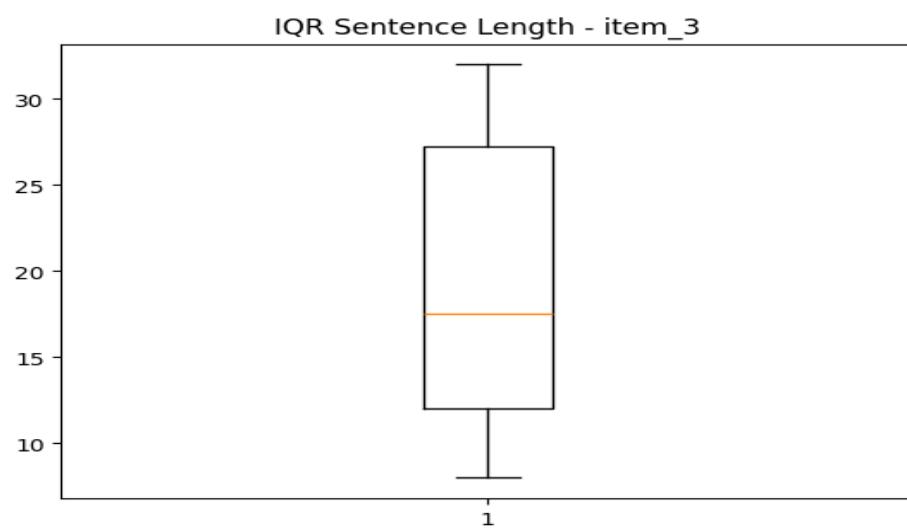
sentence_IQR_item_28.png



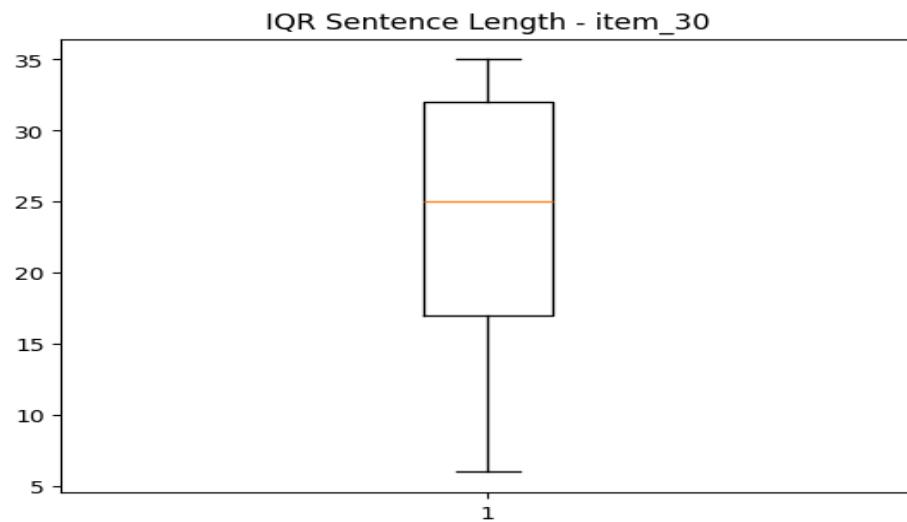
sentence_IQR_item_29.png



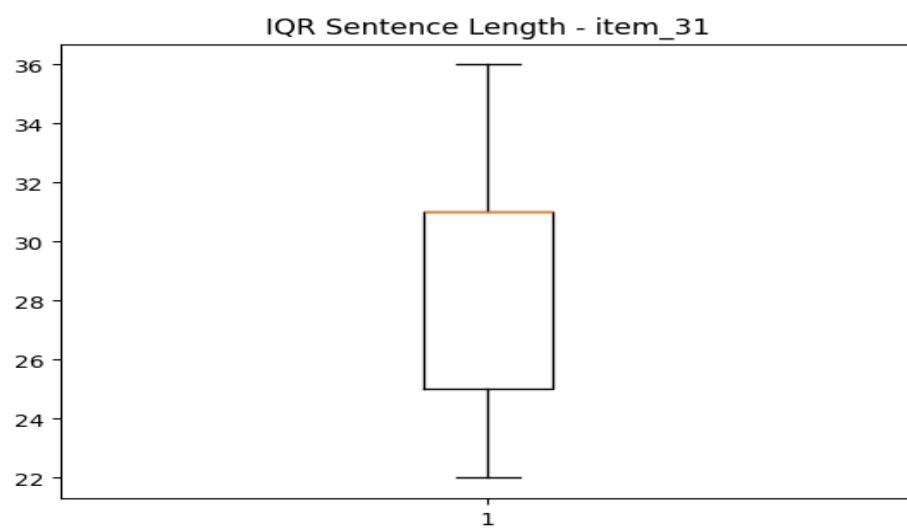
sentence_IQR_item_3.png



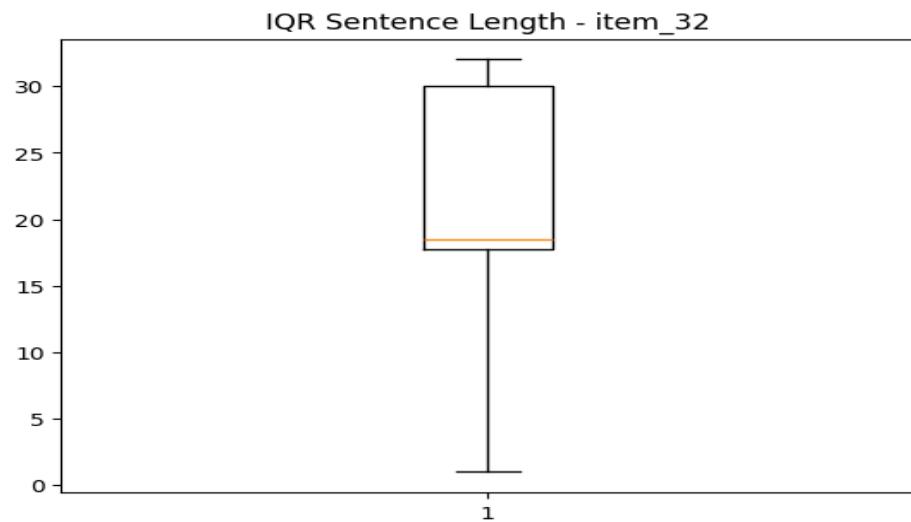
sentence_IQR_item_30.png



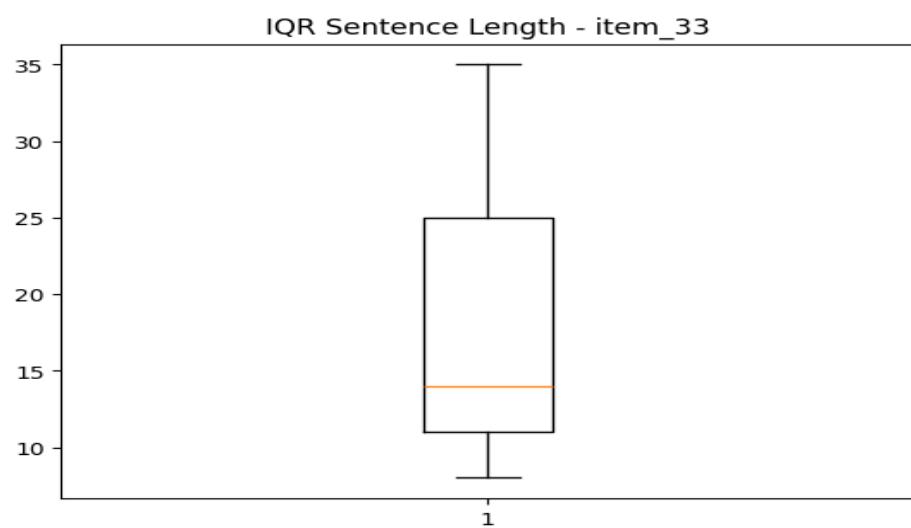
sentence_IQR_item_31.png



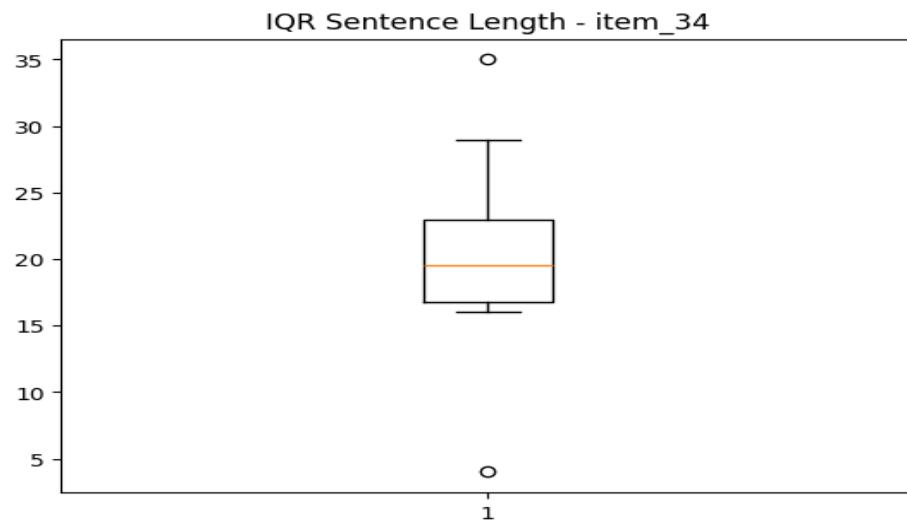
sentence_IQR_item_32.png



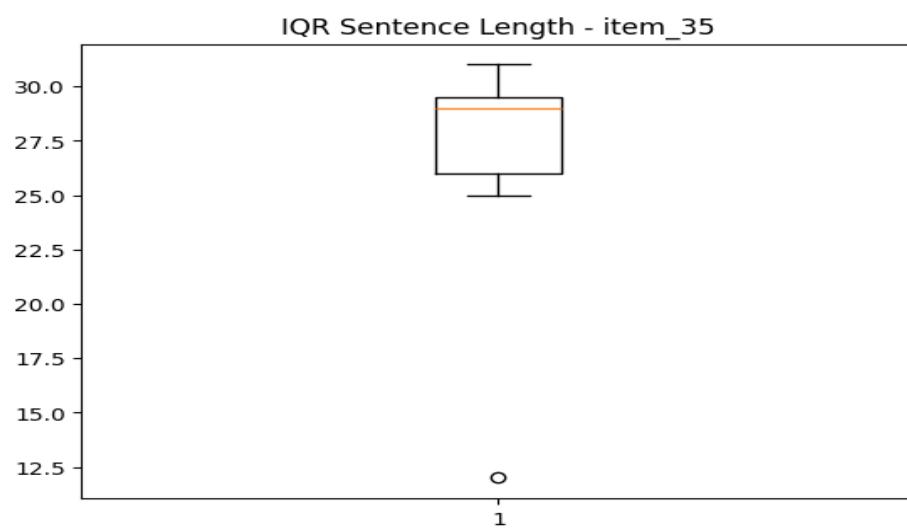
sentence_IQR_item_33.png



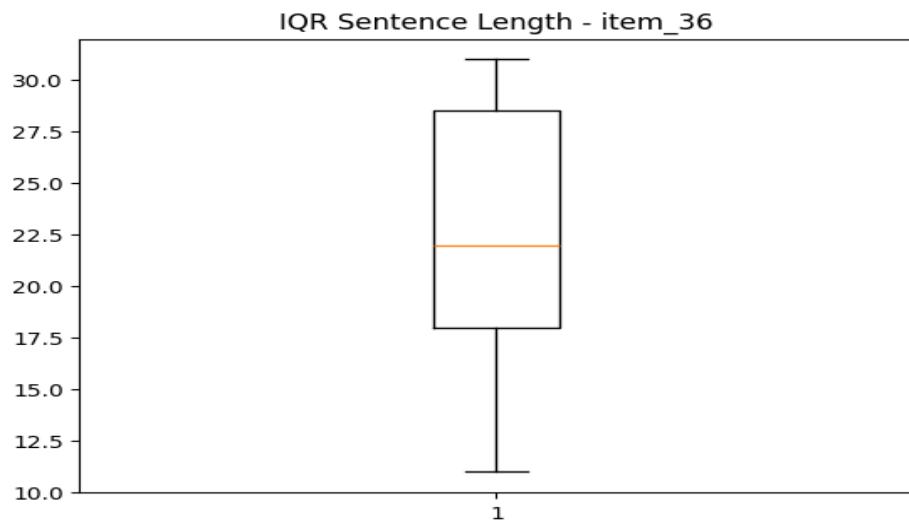
sentence_IQR_item_34.png



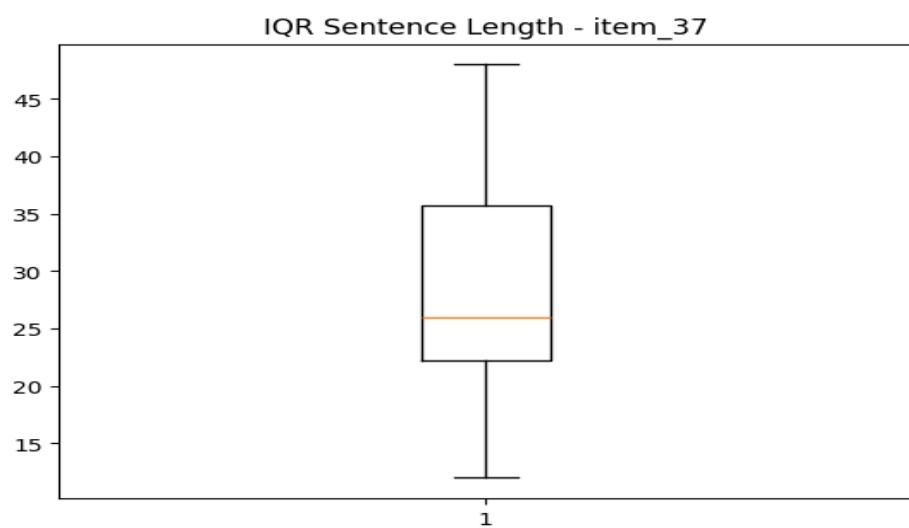
sentence_IQR_item_35.png



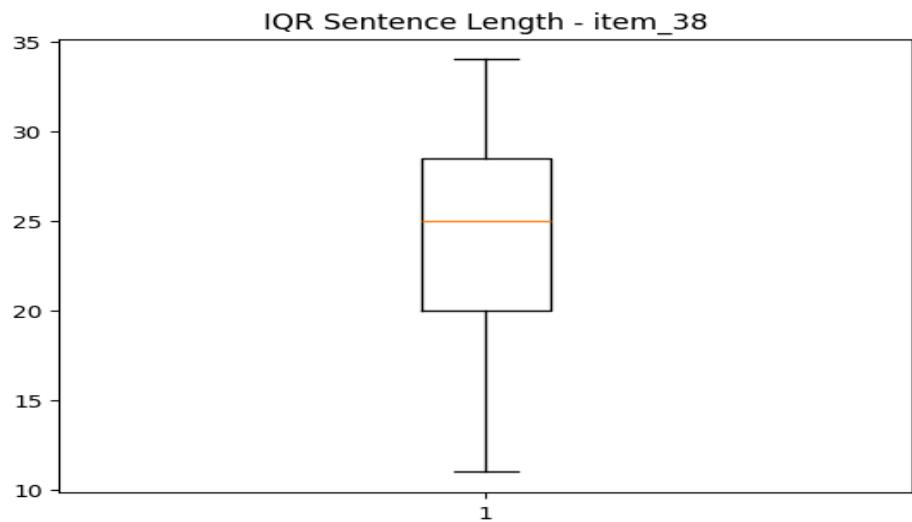
sentence_IQR_item_36.png



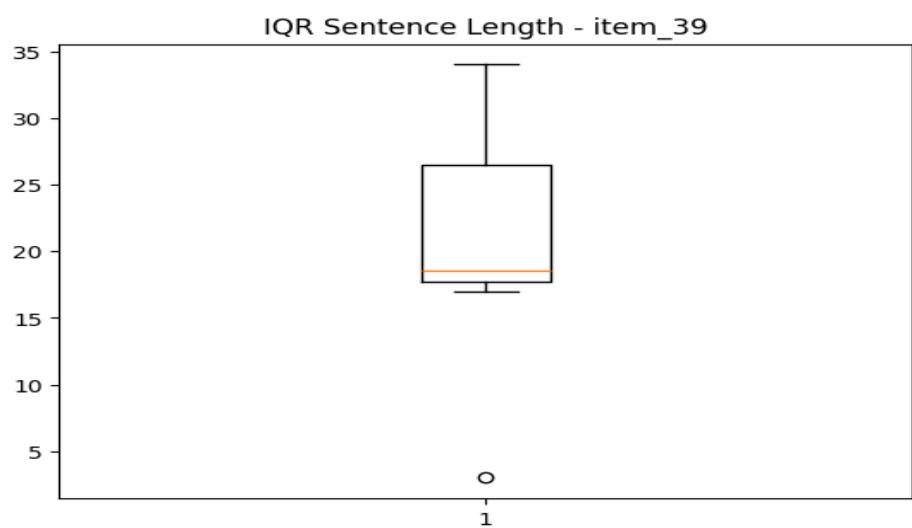
sentence_IQR_item_37.png



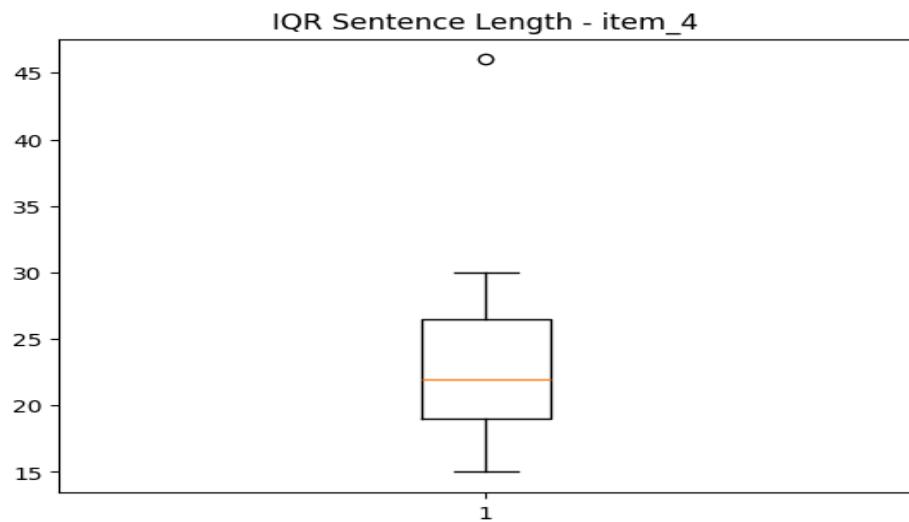
sentence_IQR_item_38.png



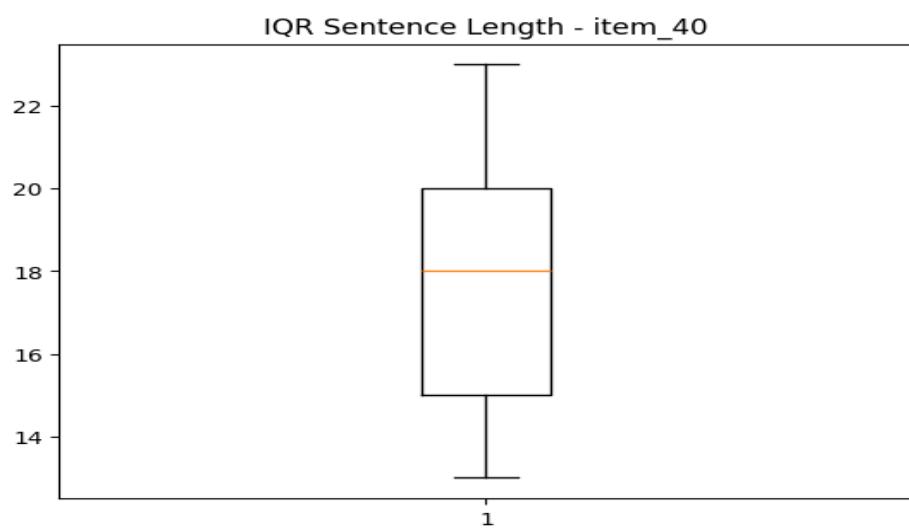
sentence_IQR_item_39.png



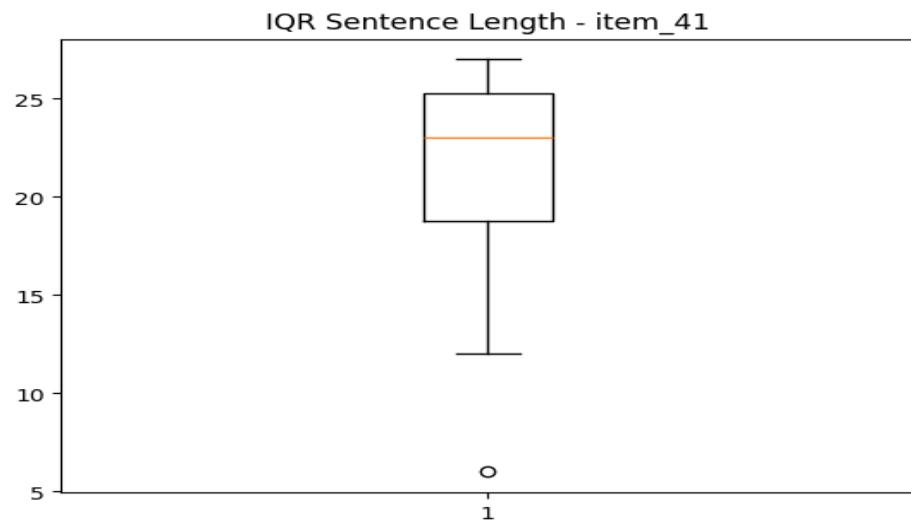
sentence_IQR_item_4.png



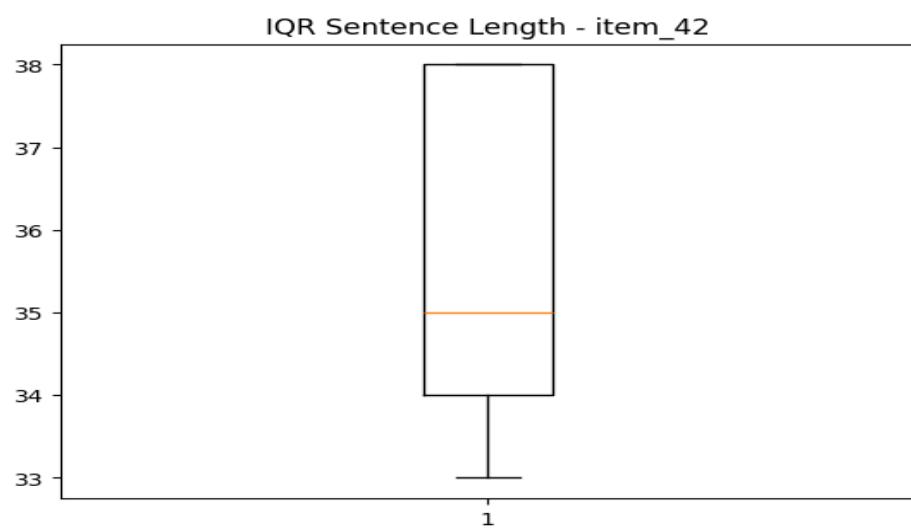
sentence_IQR_item_40.png



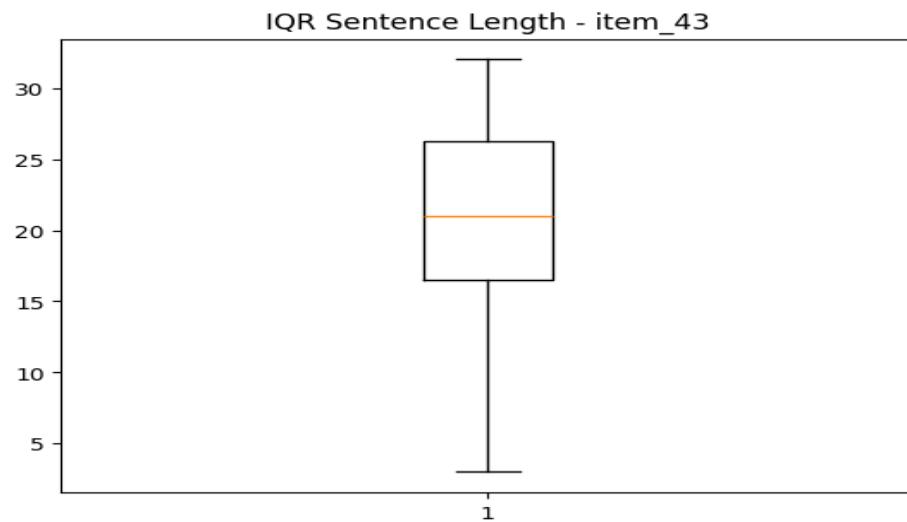
sentence_IQR_item_41.png



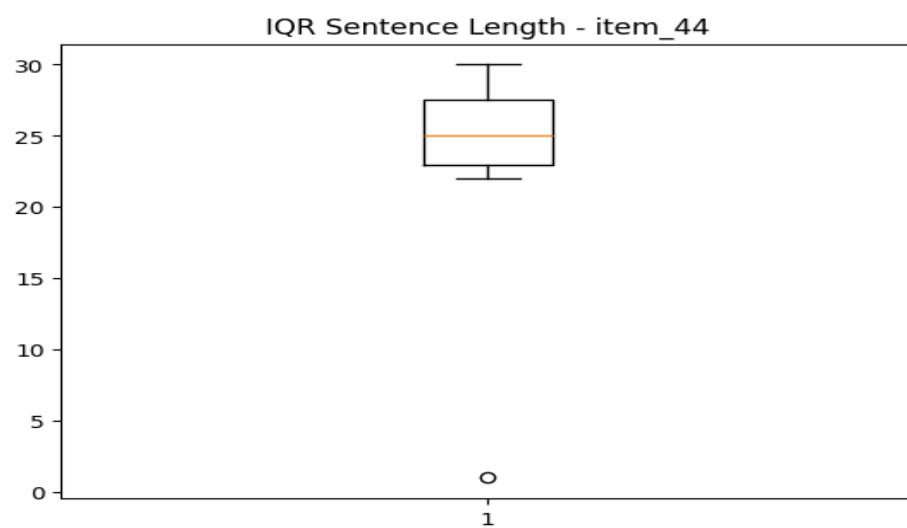
sentence_IQR_item_42.png



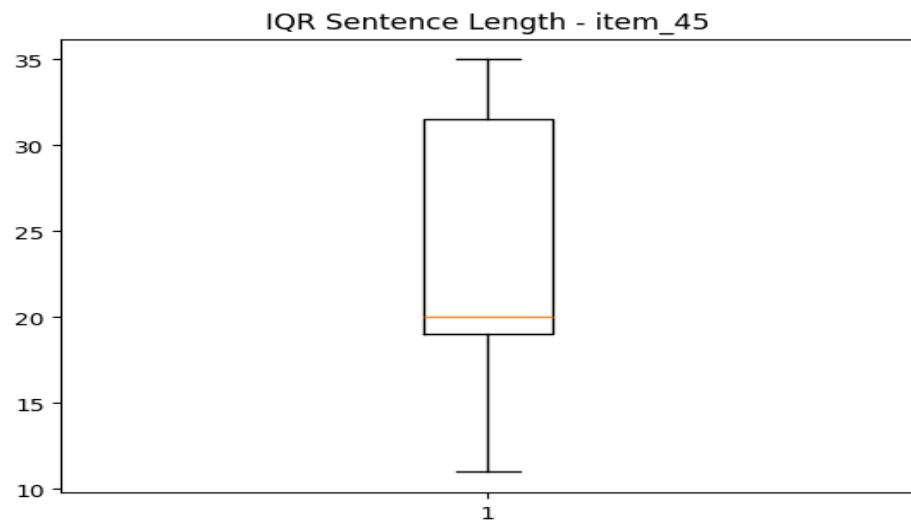
sentence_IQR_item_43.png



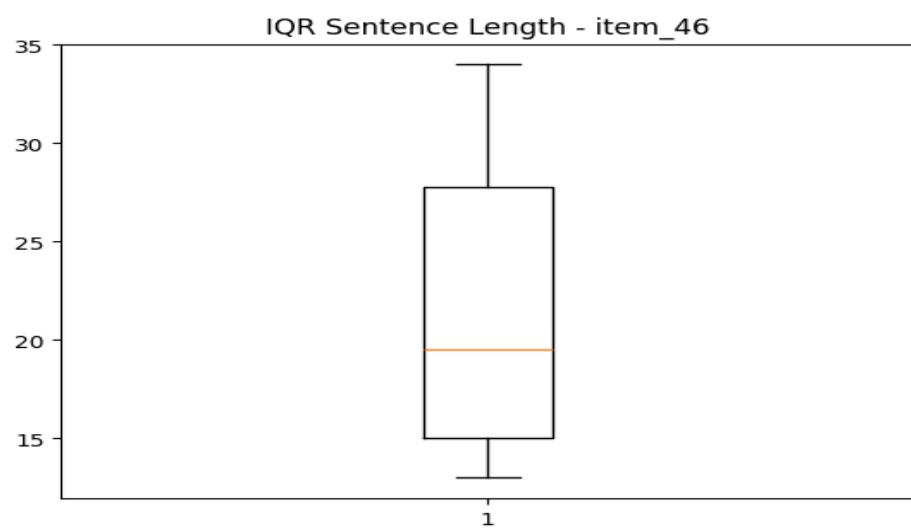
sentence_IQR_item_44.png



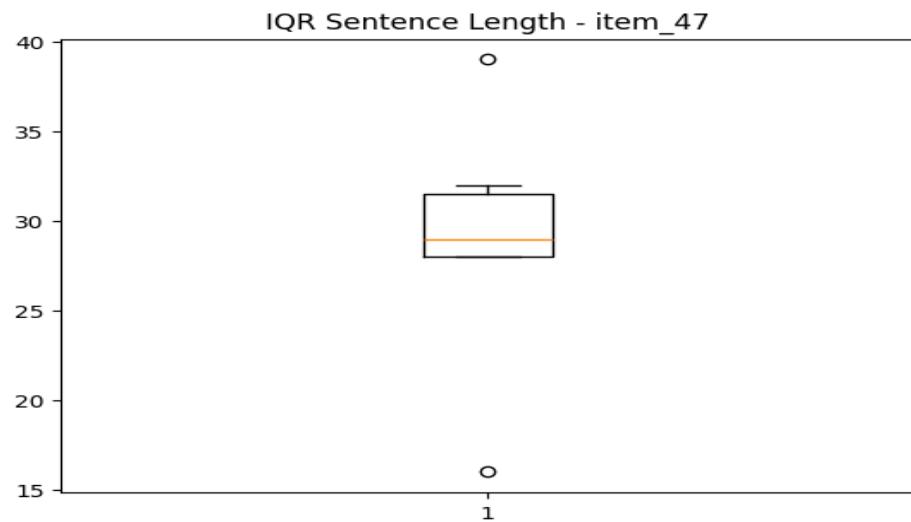
sentence_IQR_item_45.png



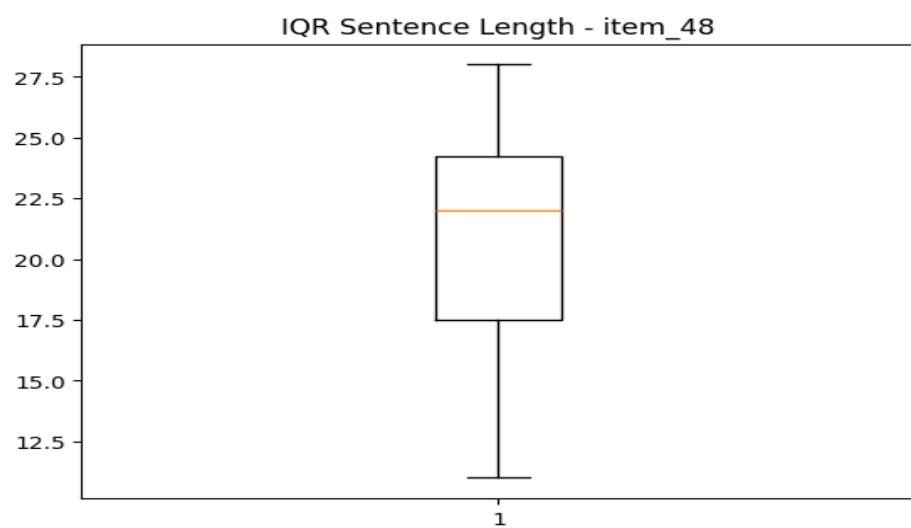
sentence_IQR_item_46.png



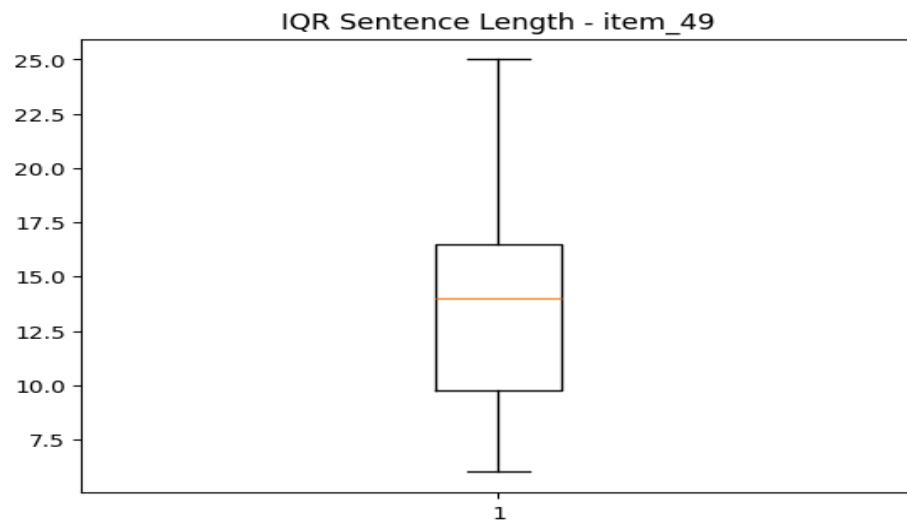
sentence_IQR_item_47.png



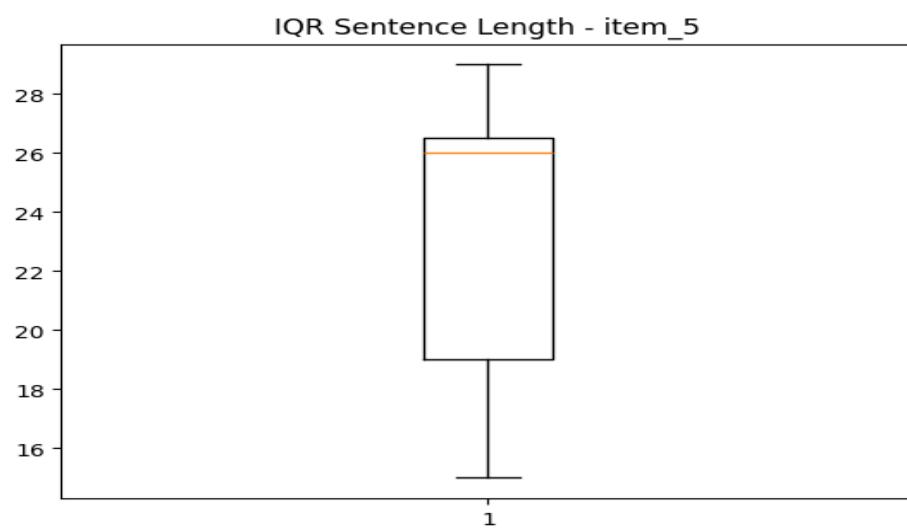
sentence_IQR_item_48.png



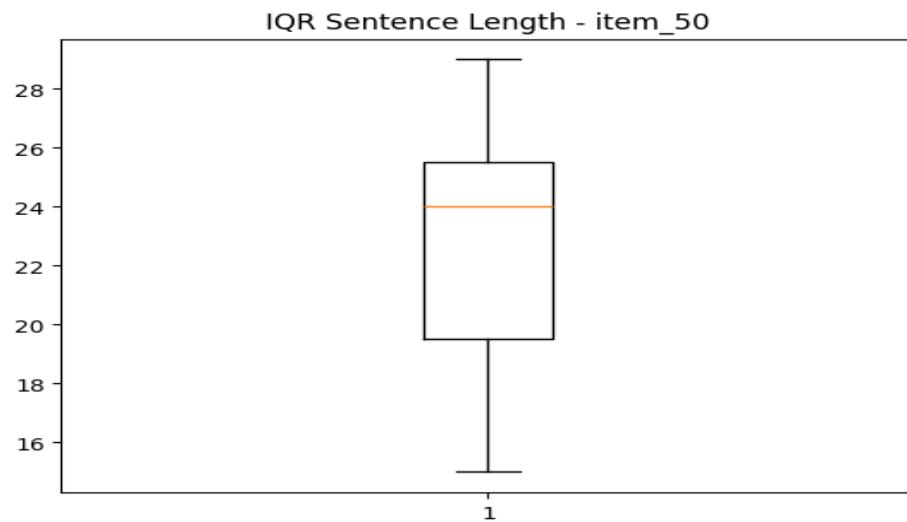
sentence_IQR_item_49.png



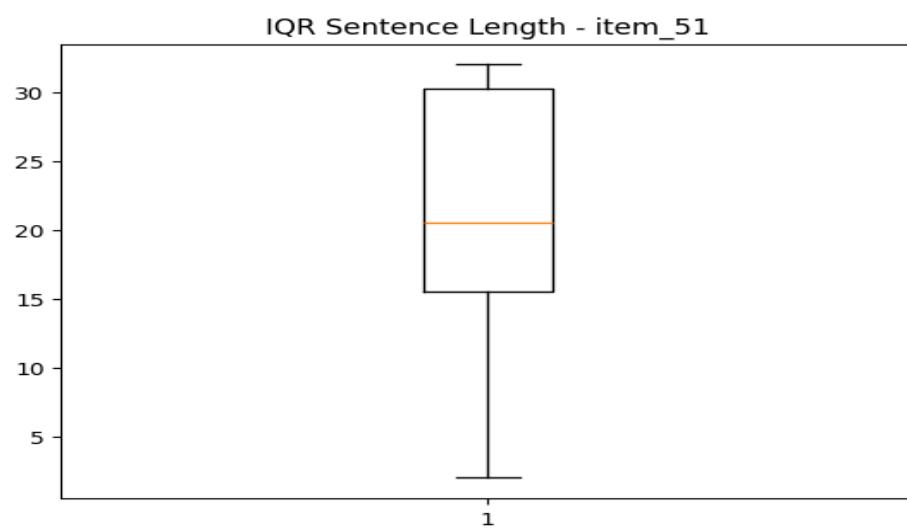
sentence_IQR_item_5.png



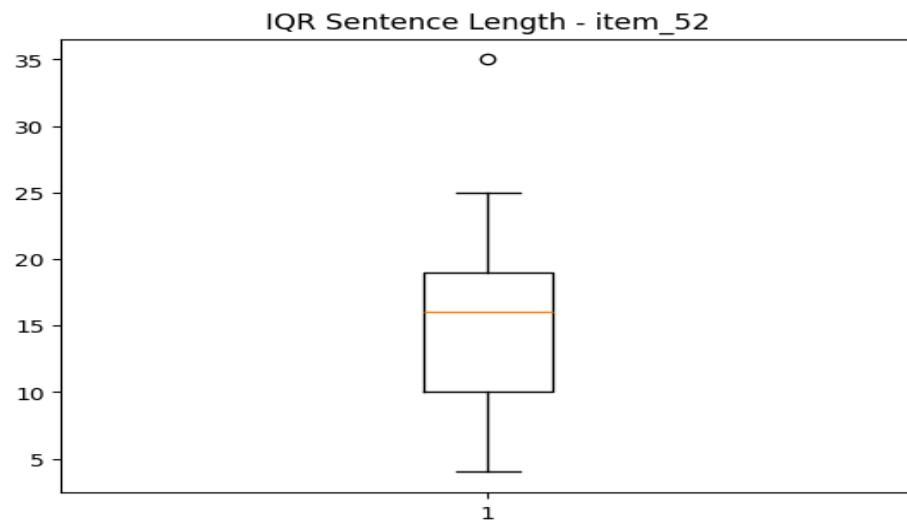
sentence_IQR_item_50.png



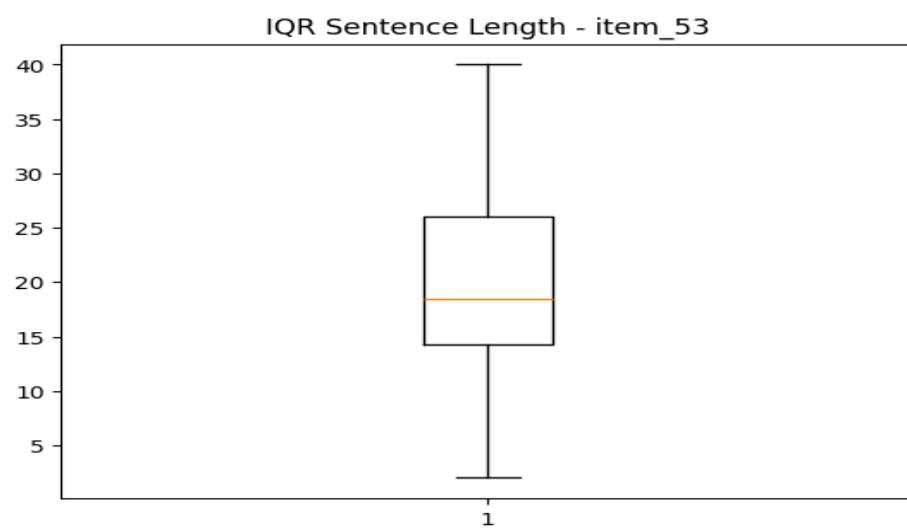
sentence_IQR_item_51.png



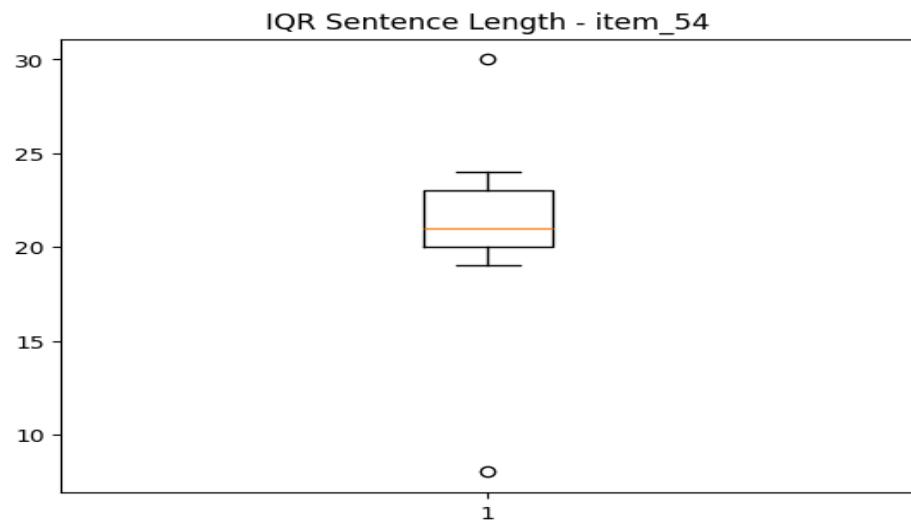
sentence_IQR_item_52.png



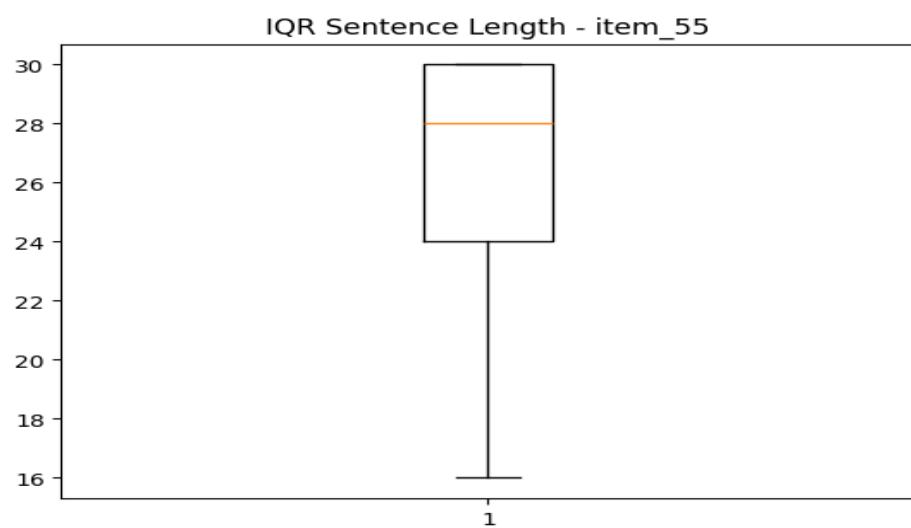
sentence_IQR_item_53.png



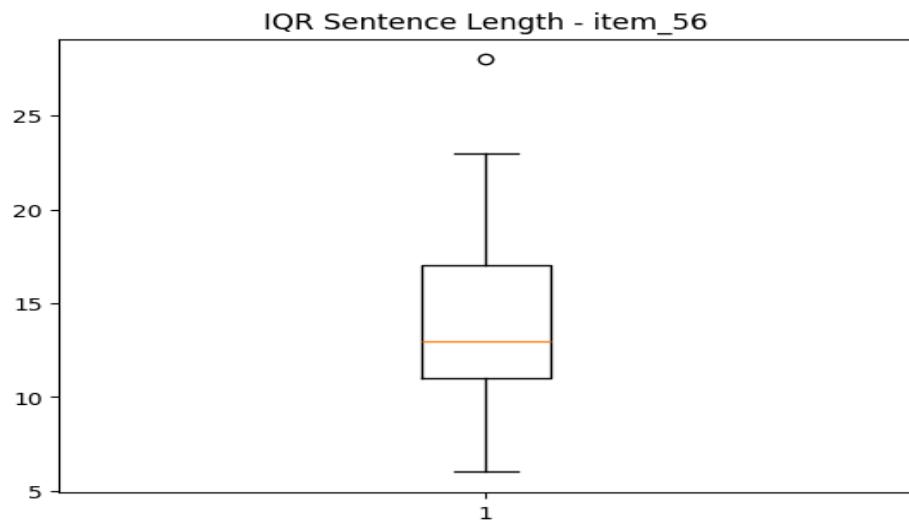
sentence_IQR_item_54.png



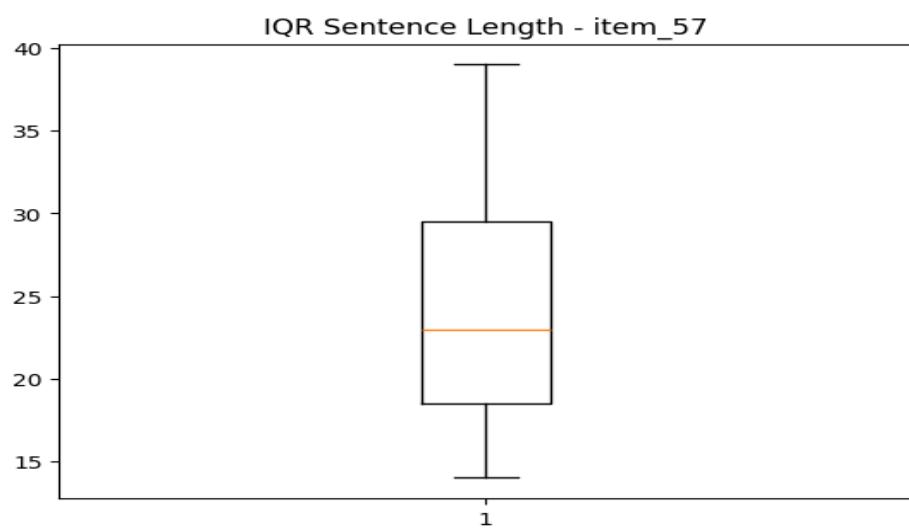
sentence_IQR_item_55.png



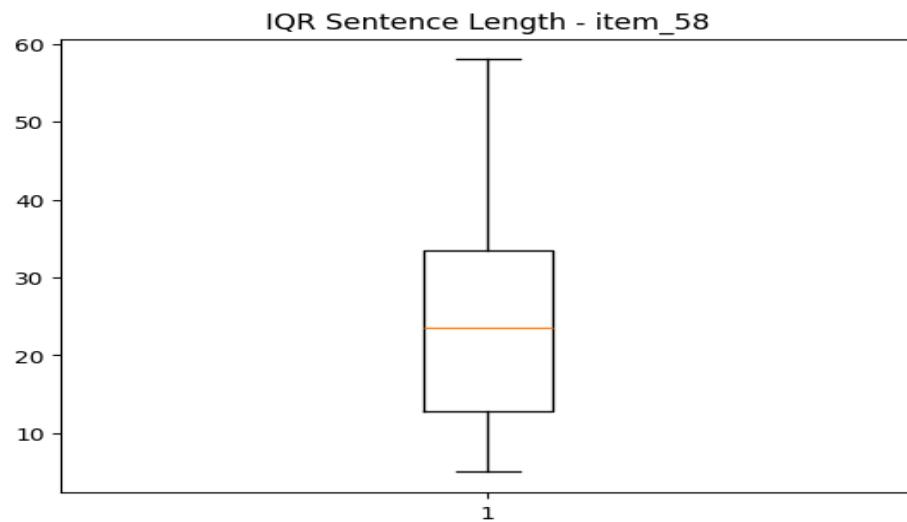
sentence_IQR_item_56.png



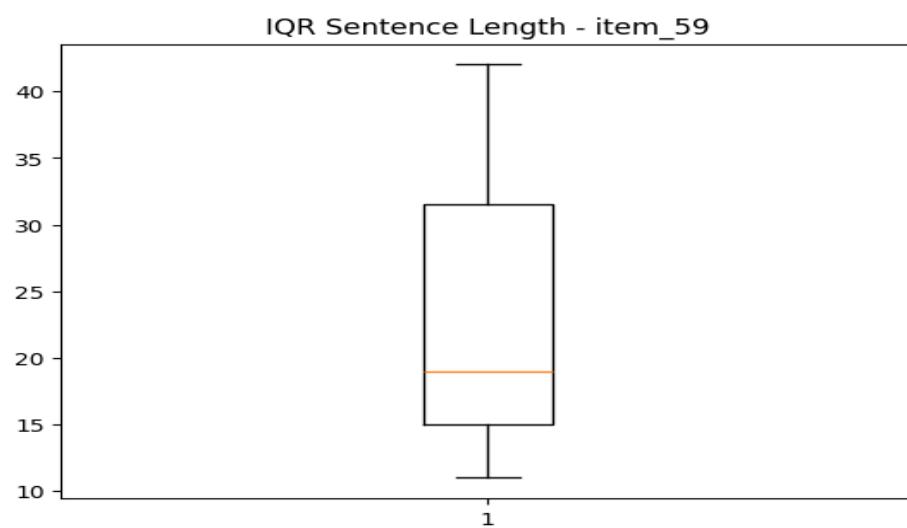
sentence_IQR_item_57.png



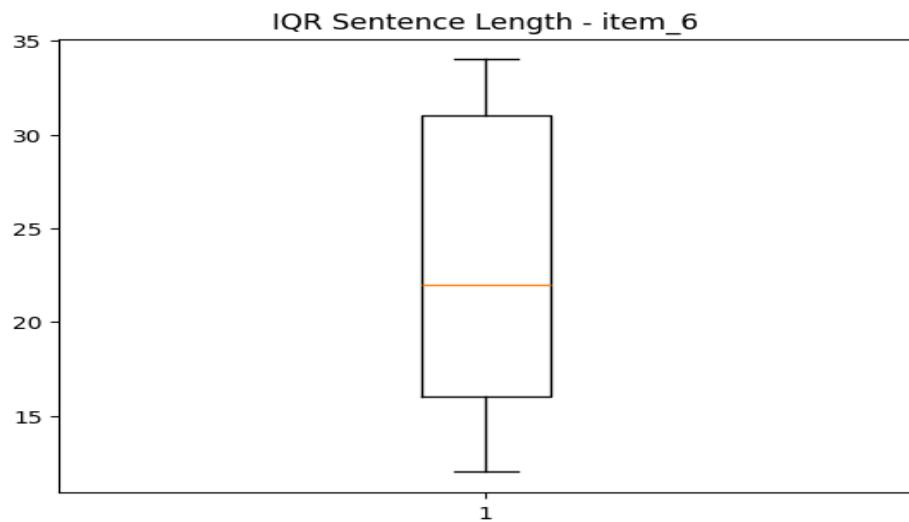
sentence_IQR_item_58.png



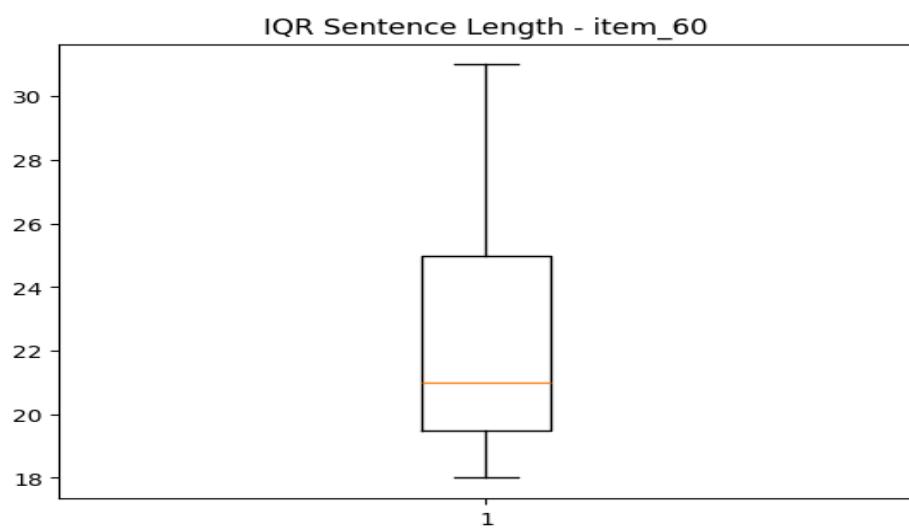
sentence_IQR_item_59.png



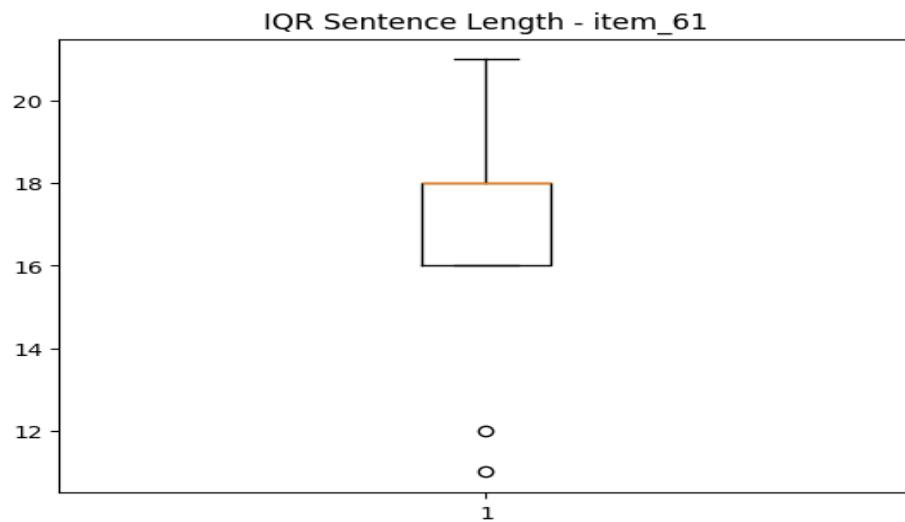
sentence_IQR_item_6.png



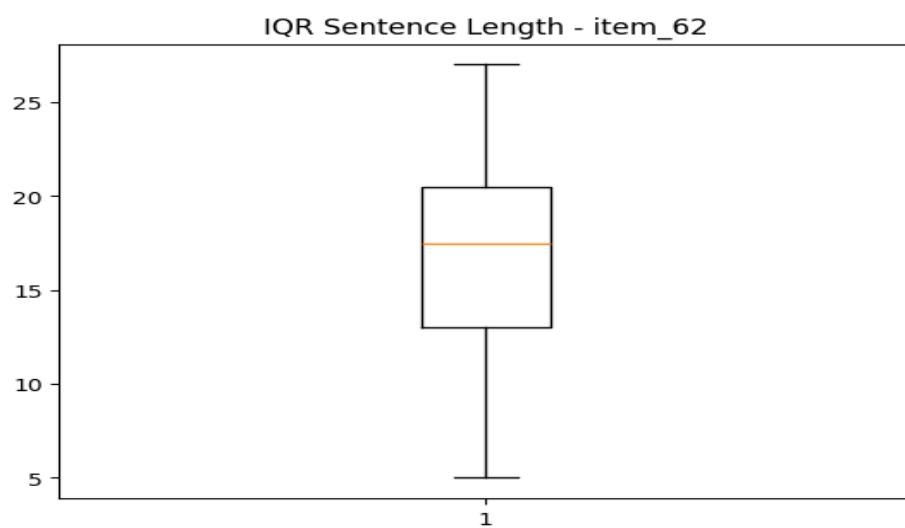
sentence_IQR_item_60.png



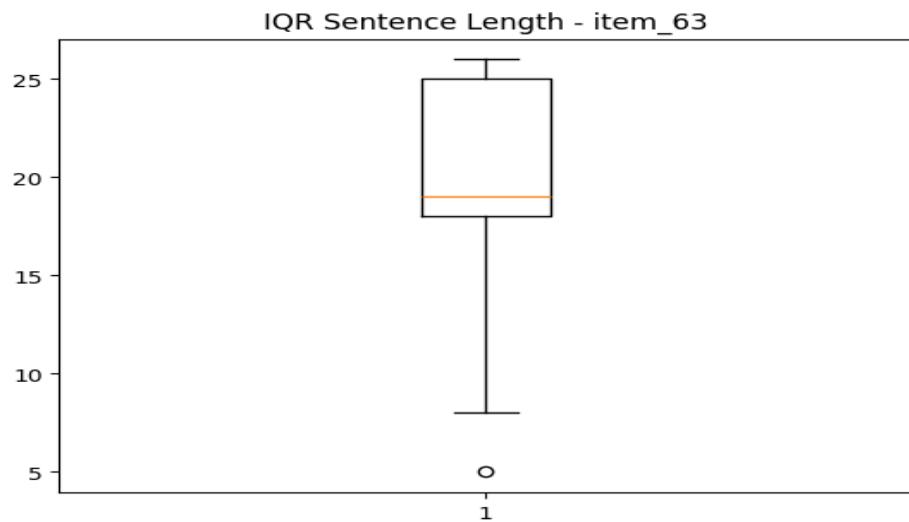
sentence_IQR_item_61.png



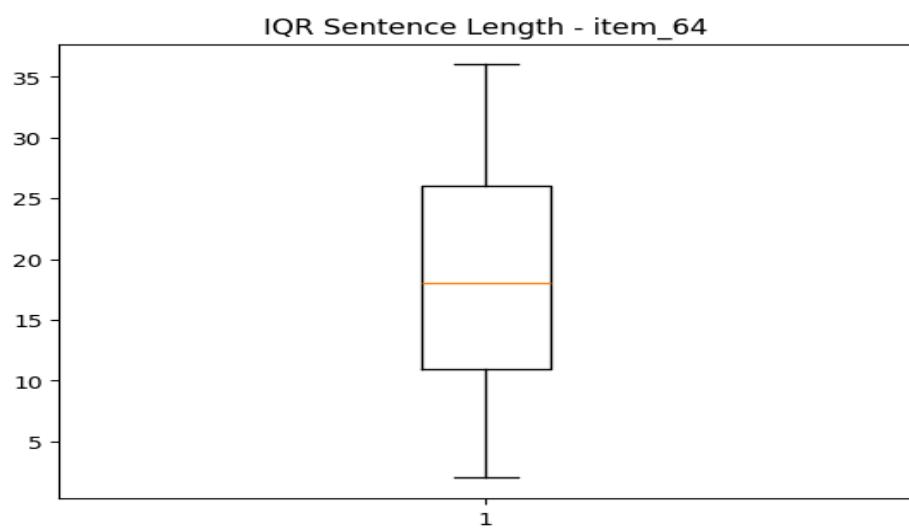
sentence_IQR_item_62.png



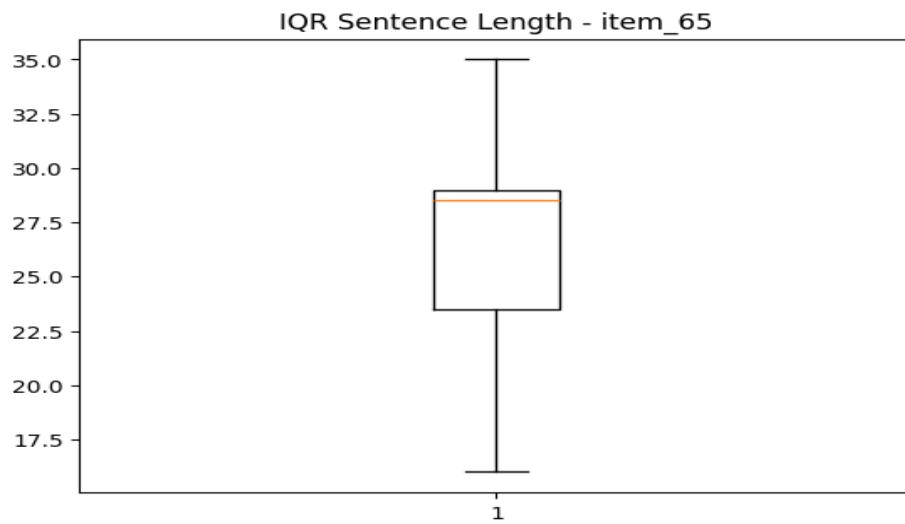
sentence_IQR_item_63.png



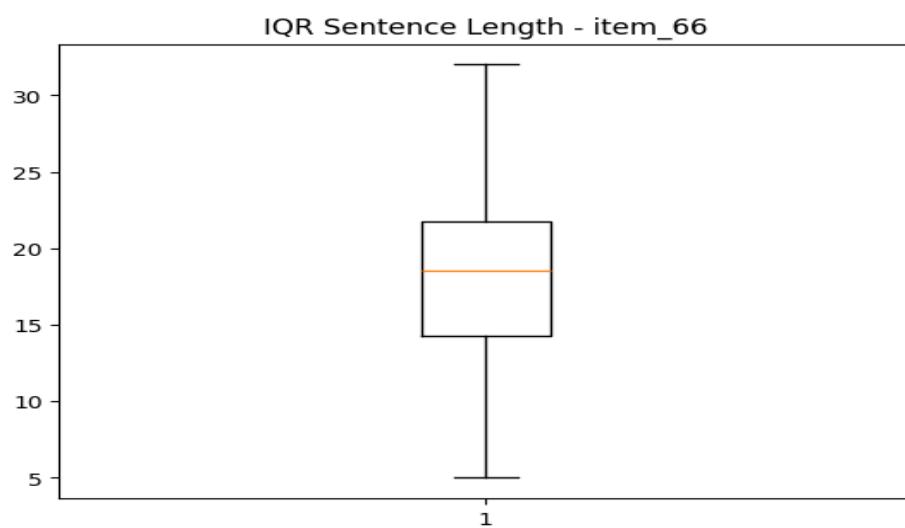
sentence_IQR_item_64.png



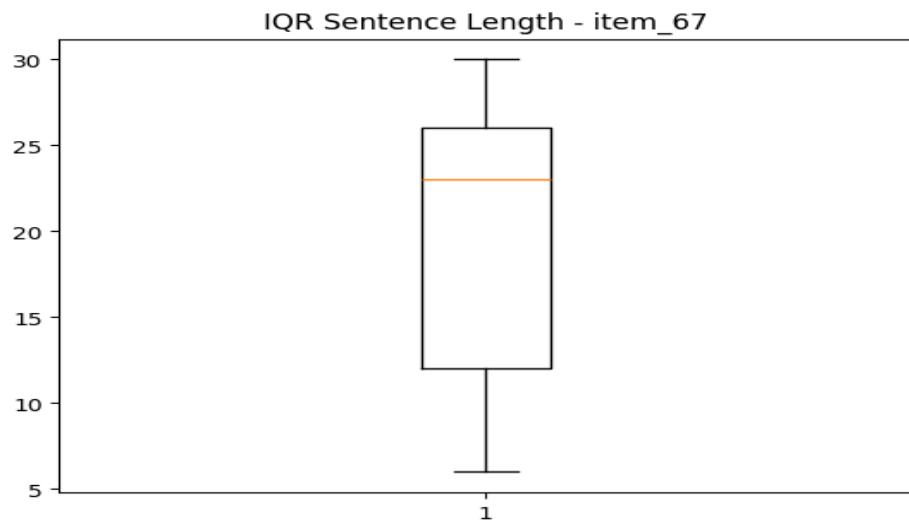
sentence_IQR_item_65.png



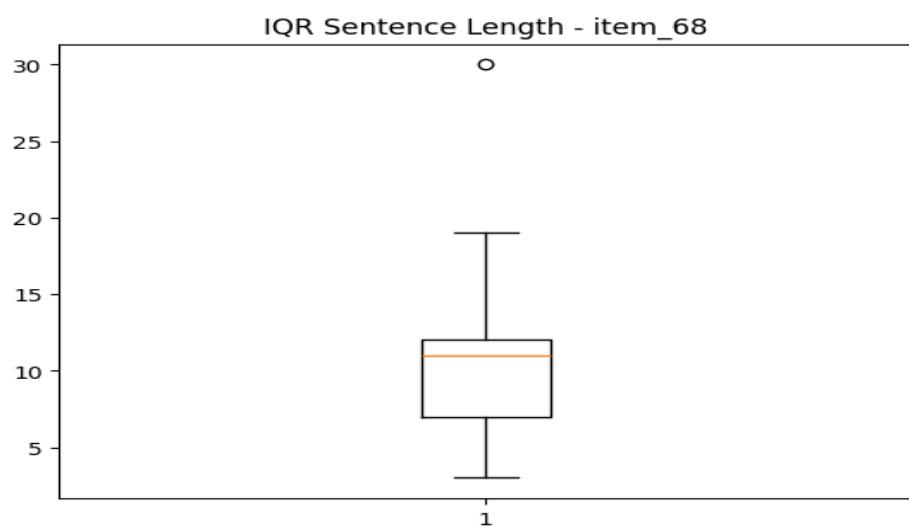
sentence_IQR_item_66.png



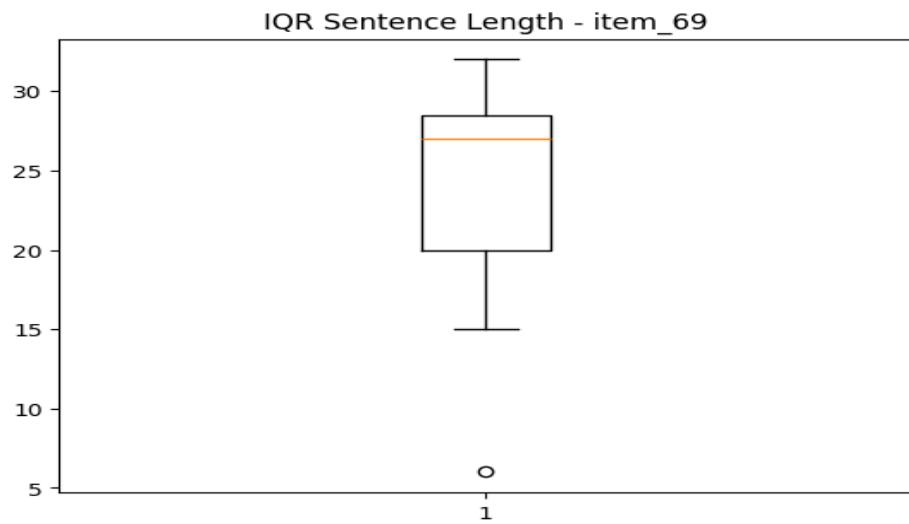
sentence_IQR_item_67.png



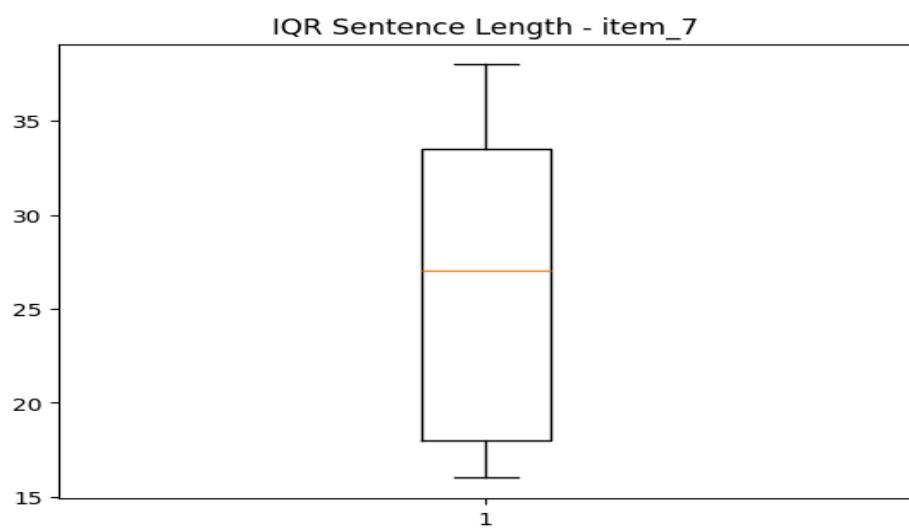
sentence_IQR_item_68.png



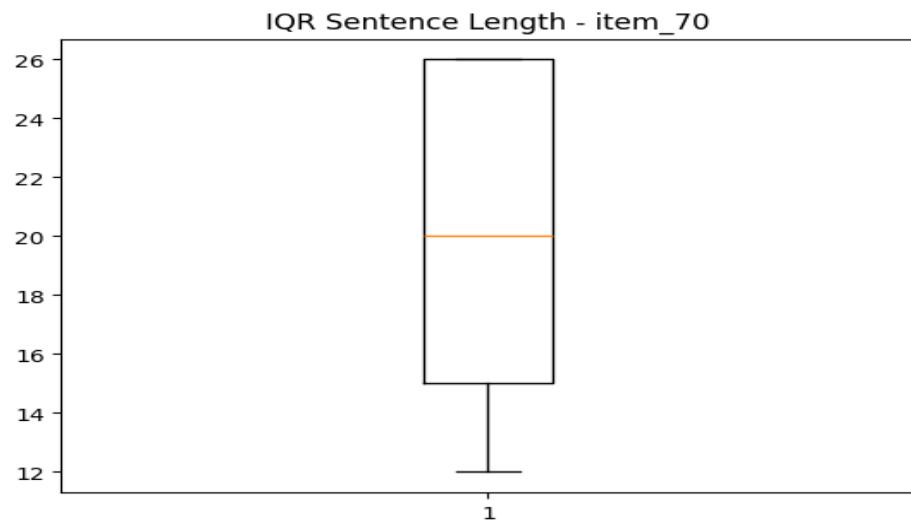
sentence_IQR_item_69.png



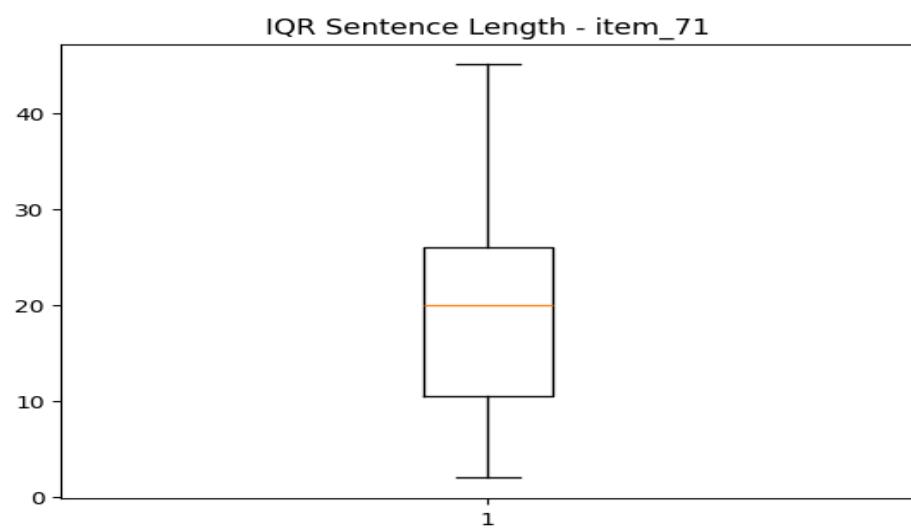
sentence_IQR_item_7.png



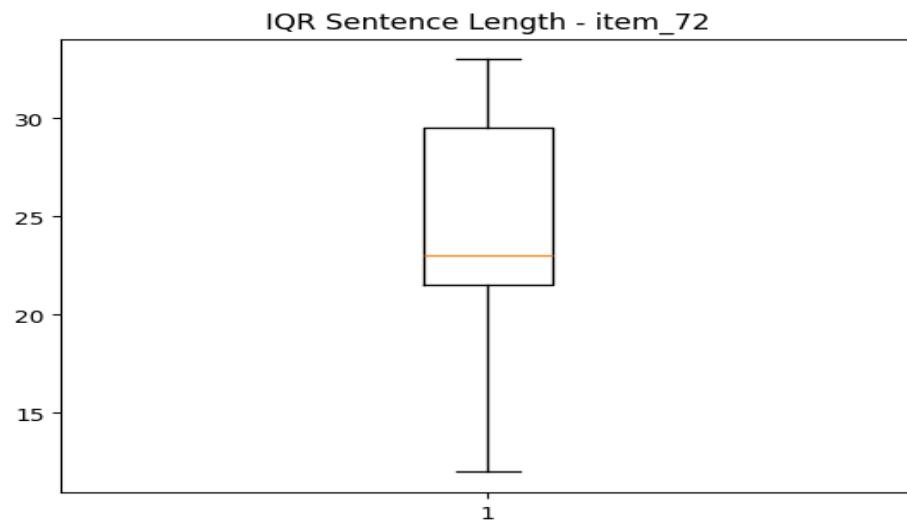
sentence_IQR_item_70.png



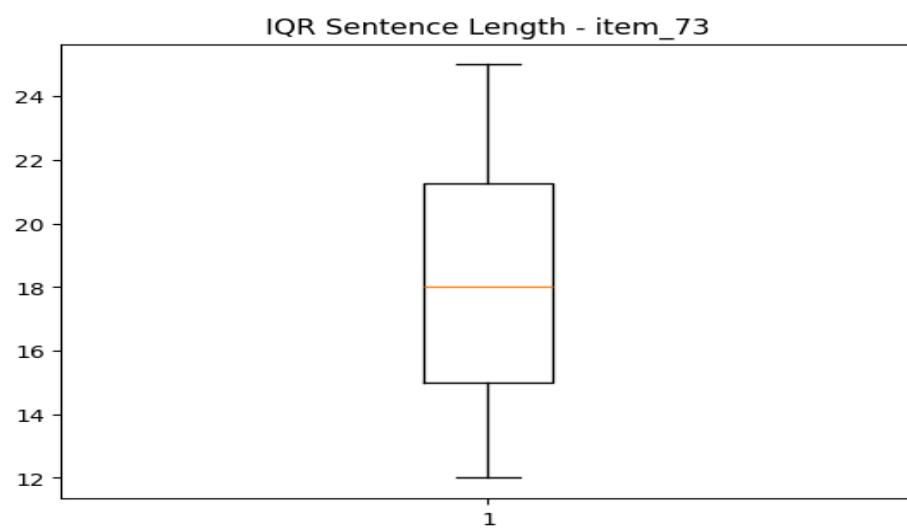
sentence_IQR_item_71.png



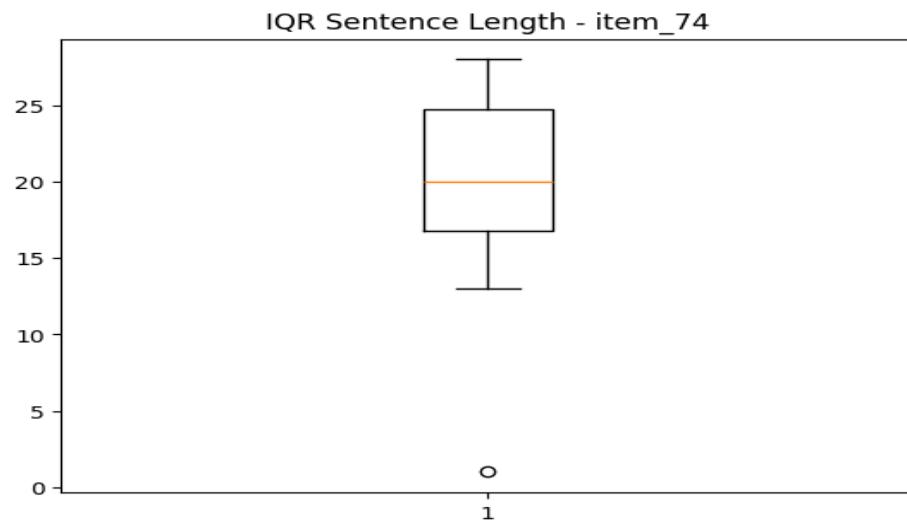
sentence_IQR_item_72.png



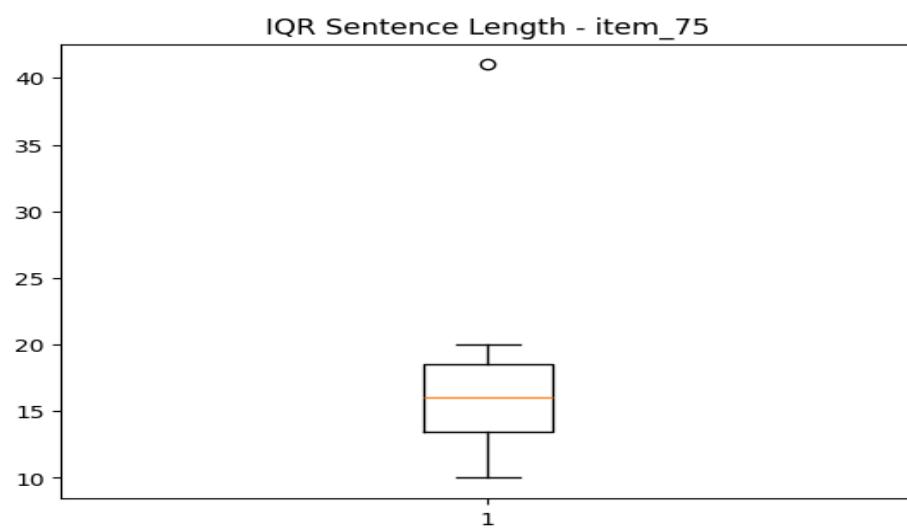
sentence_IQR_item_73.png



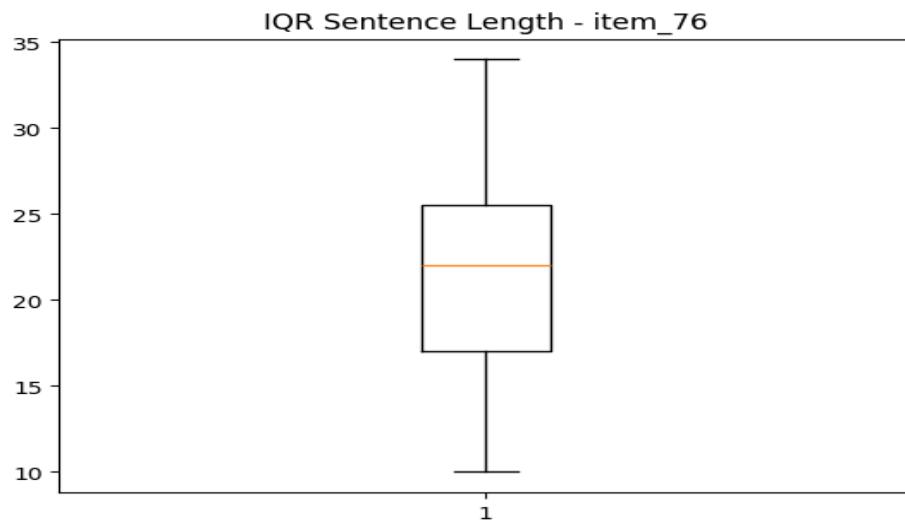
sentence_IQR_item_74.png



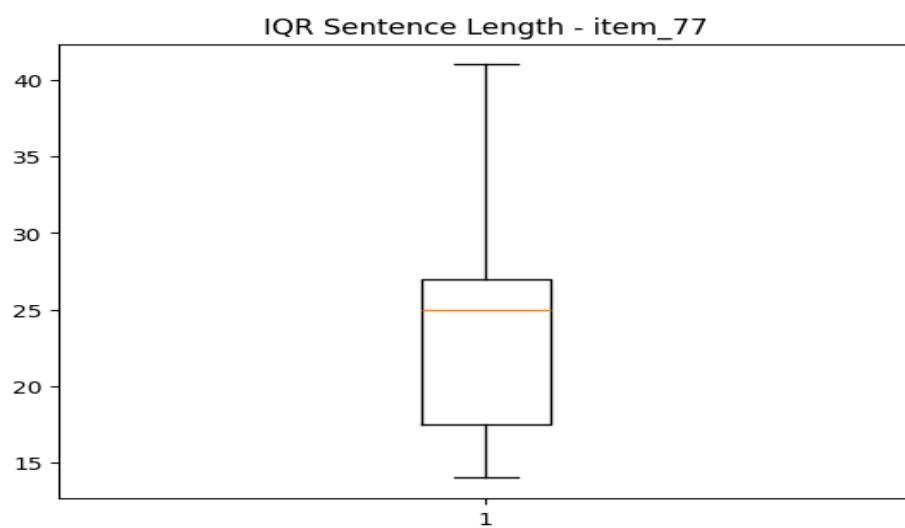
sentence_IQR_item_75.png



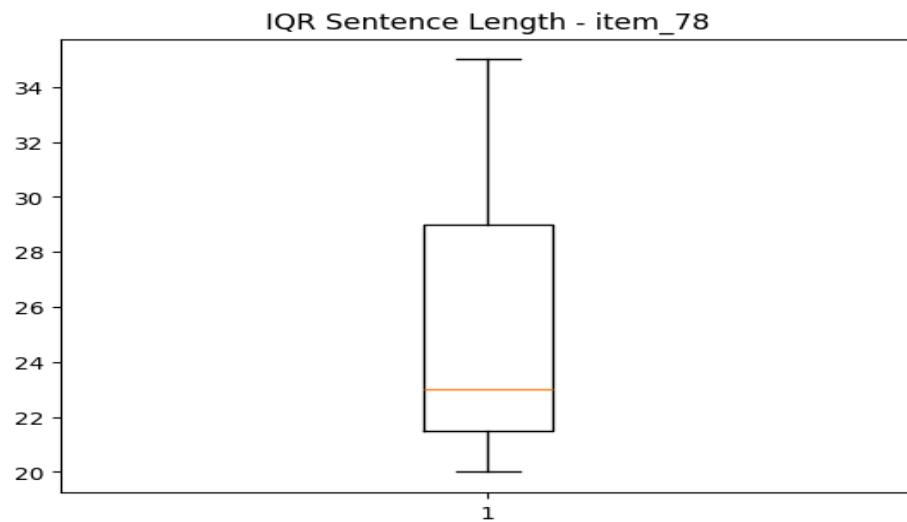
sentence_IQR_item_76.png



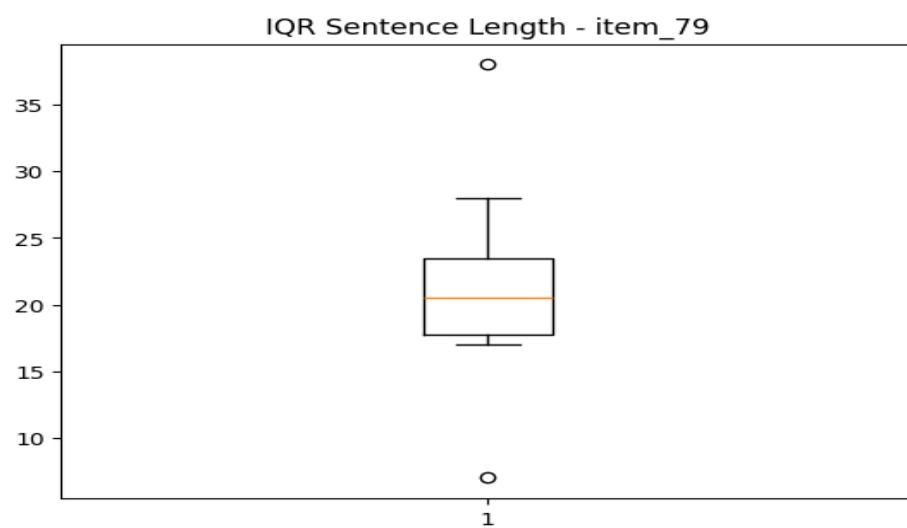
sentence_IQR_item_77.png



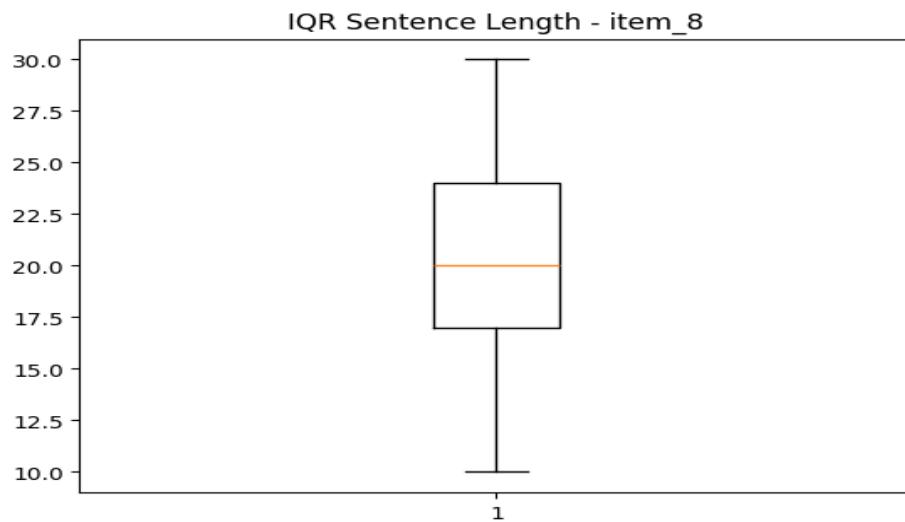
sentence_IQR_item_78.png



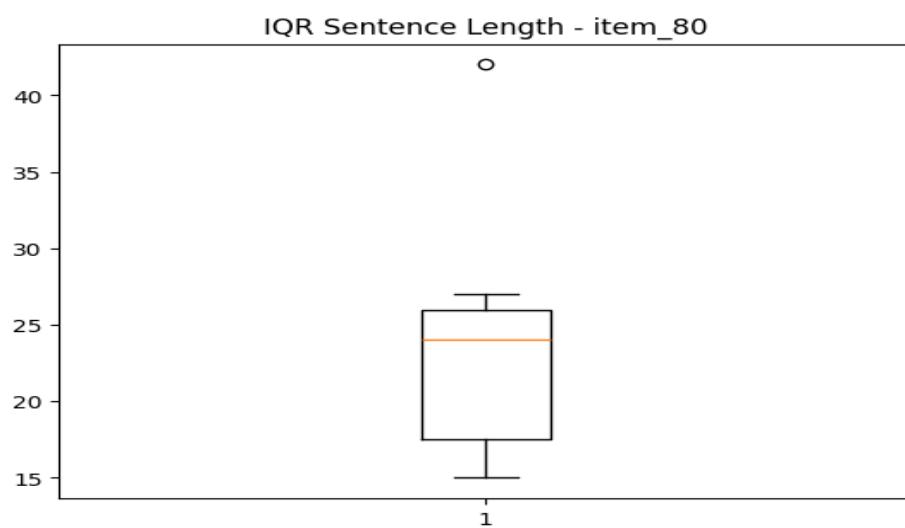
sentence_IQR_item_79.png



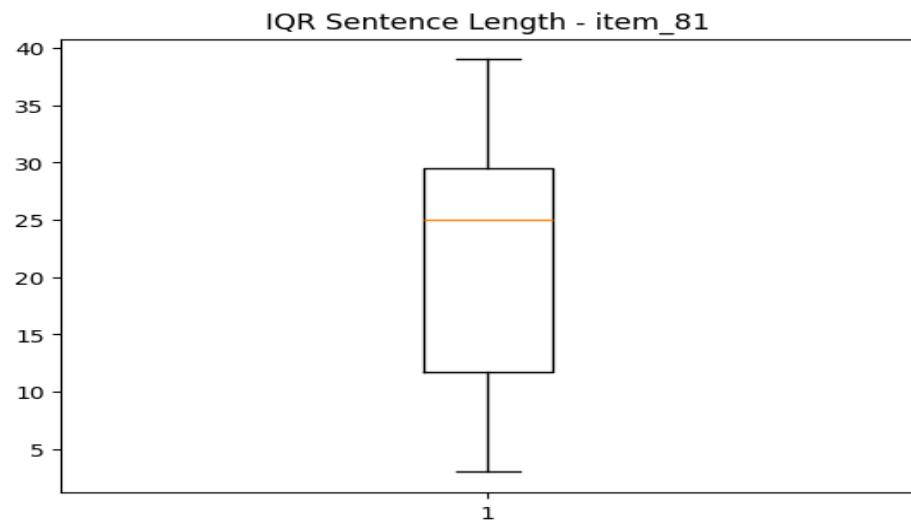
sentence_IQR_item_8.png



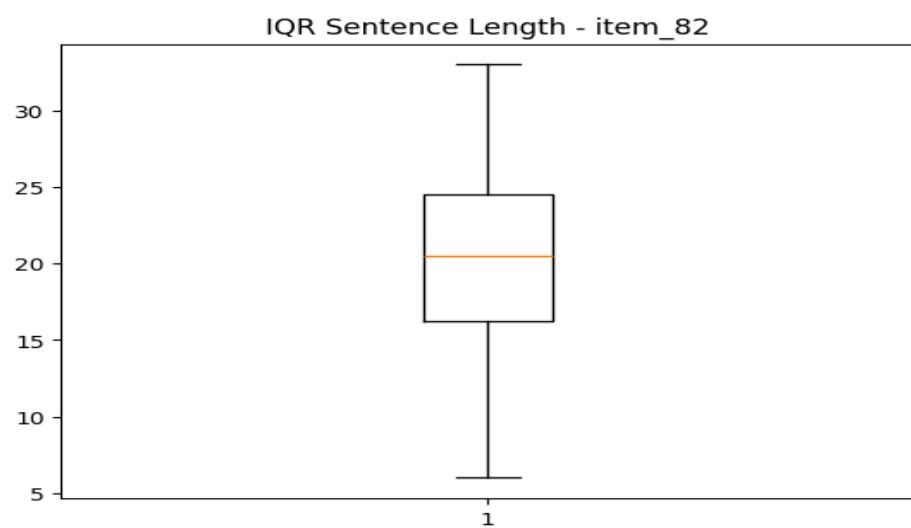
sentence_IQR_item_80.png



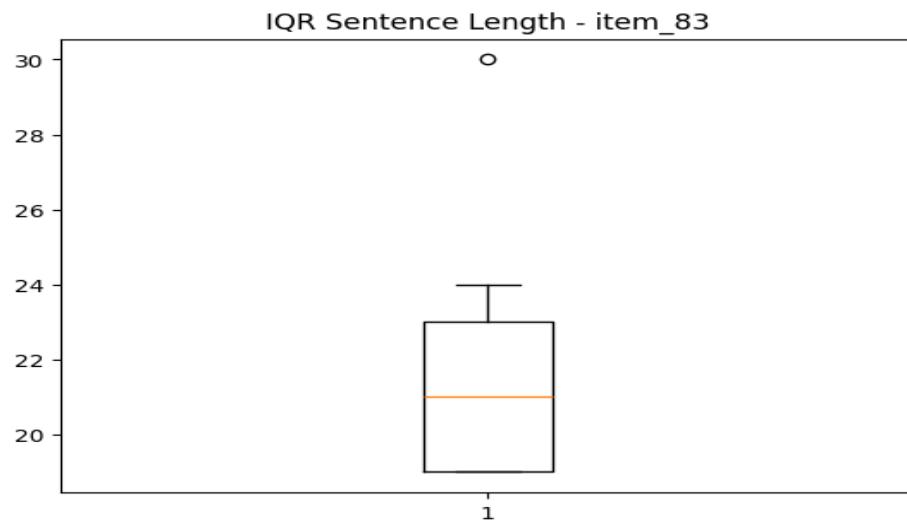
sentence_IQR_item_81.png



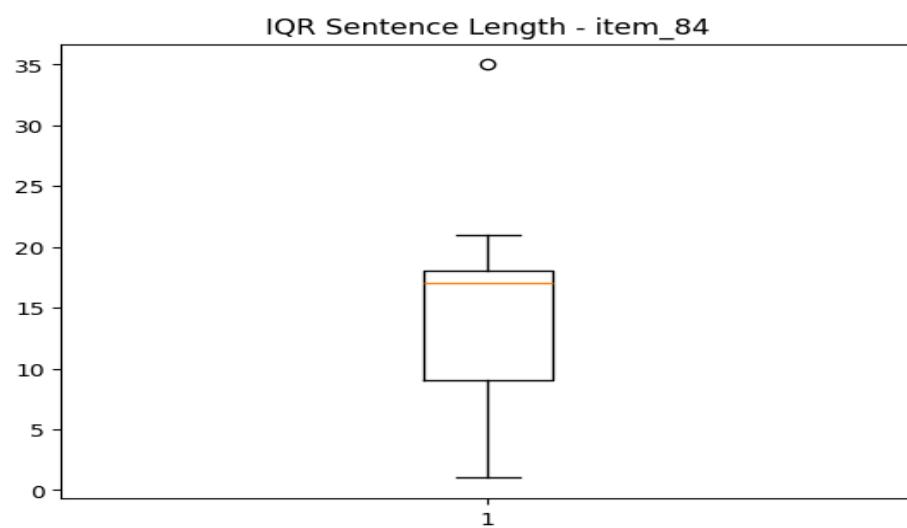
sentence_IQR_item_82.png



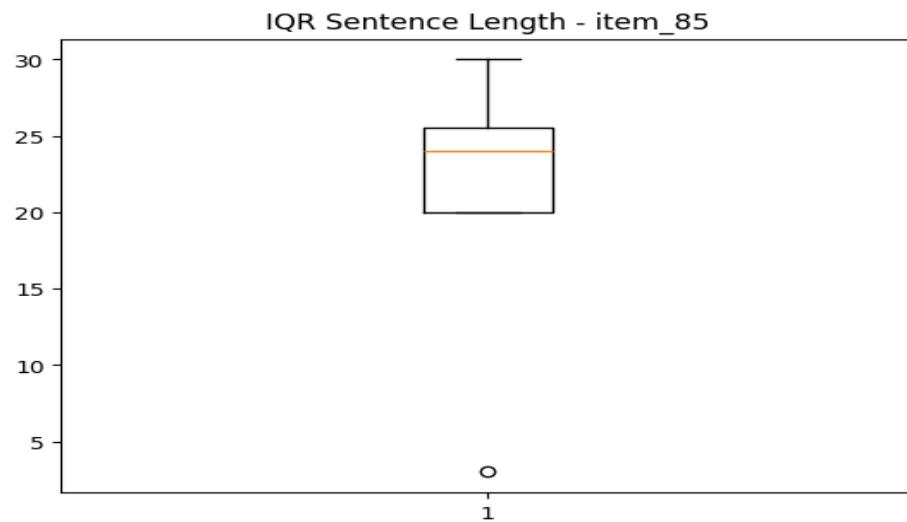
sentence_IQR_item_83.png



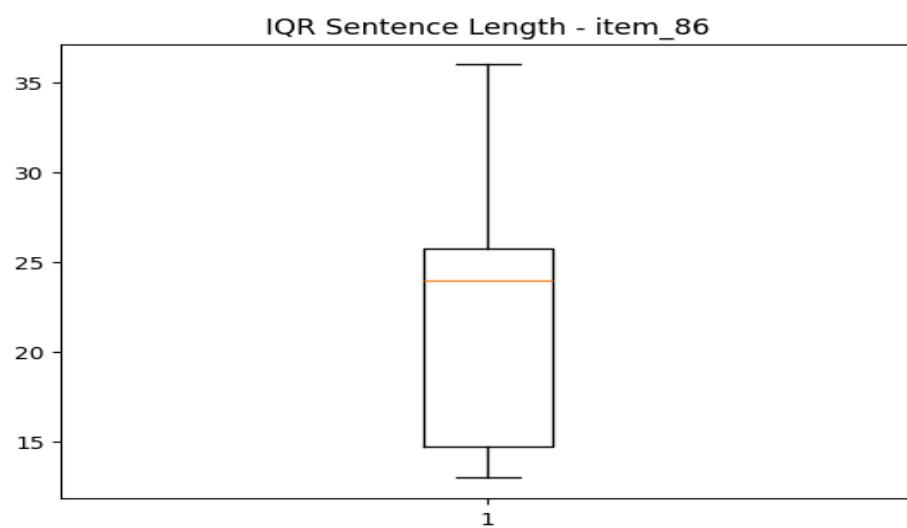
sentence_IQR_item_84.png



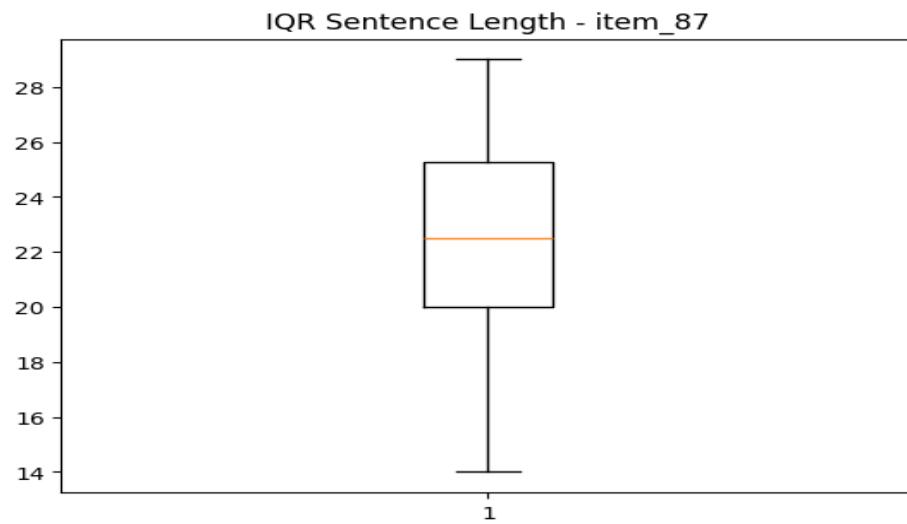
sentence_IQR_item_85.png



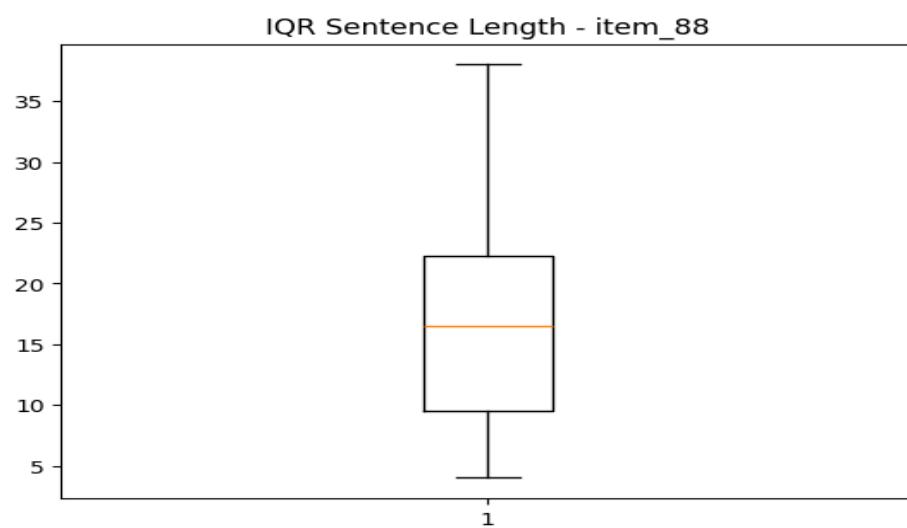
sentence_IQR_item_86.png



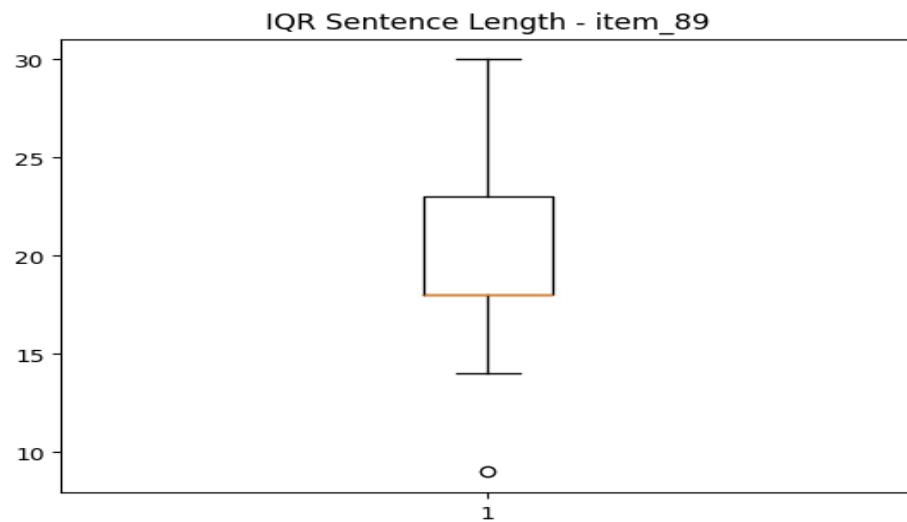
sentence_IQR_item_87.png



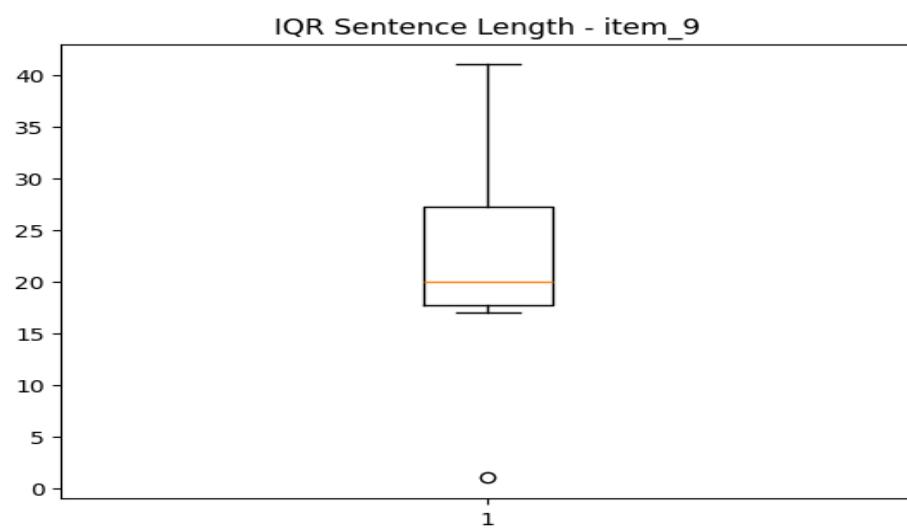
sentence_IQR_item_88.png



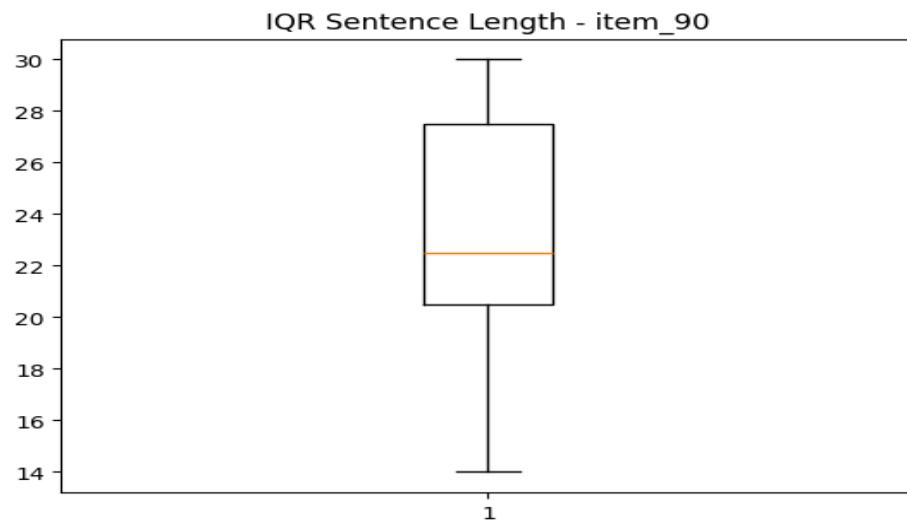
sentence_IQR_item_89.png



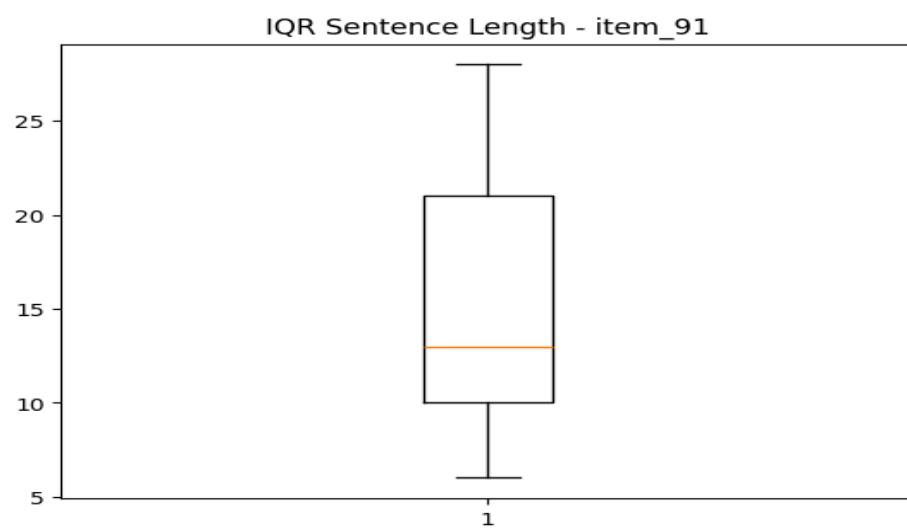
sentence_IQR_item_9.png



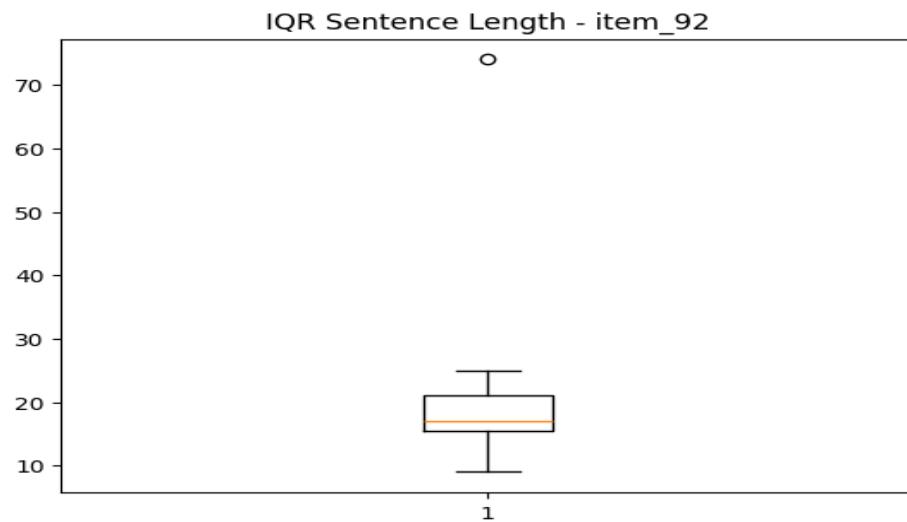
sentence_IQR_item_90.png



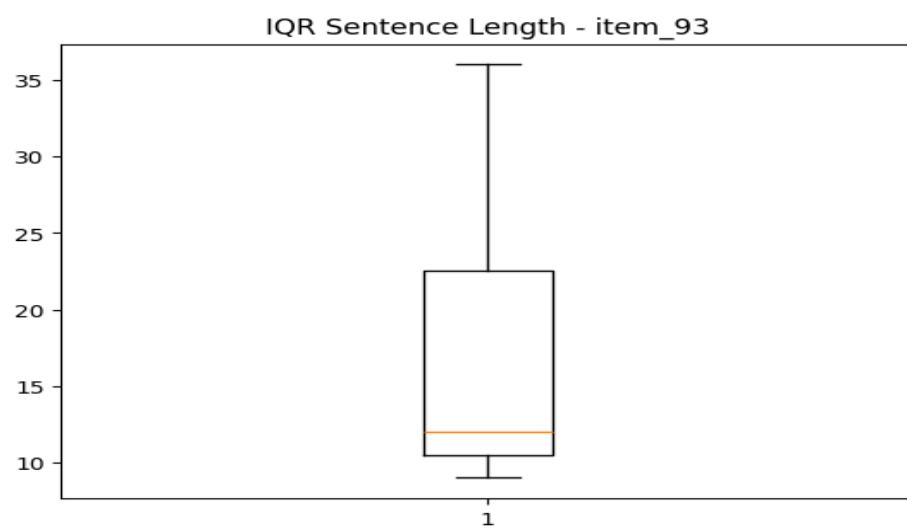
sentence_IQR_item_91.png



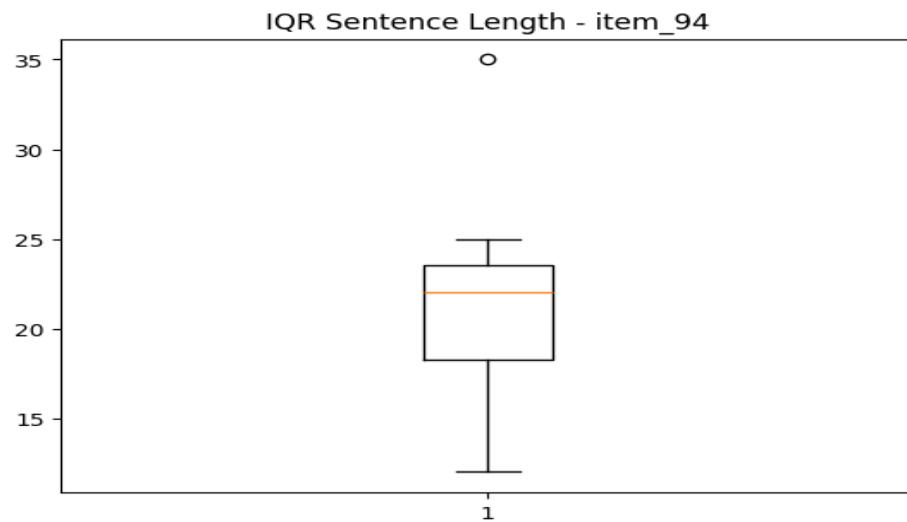
sentence_IQR_item_92.png



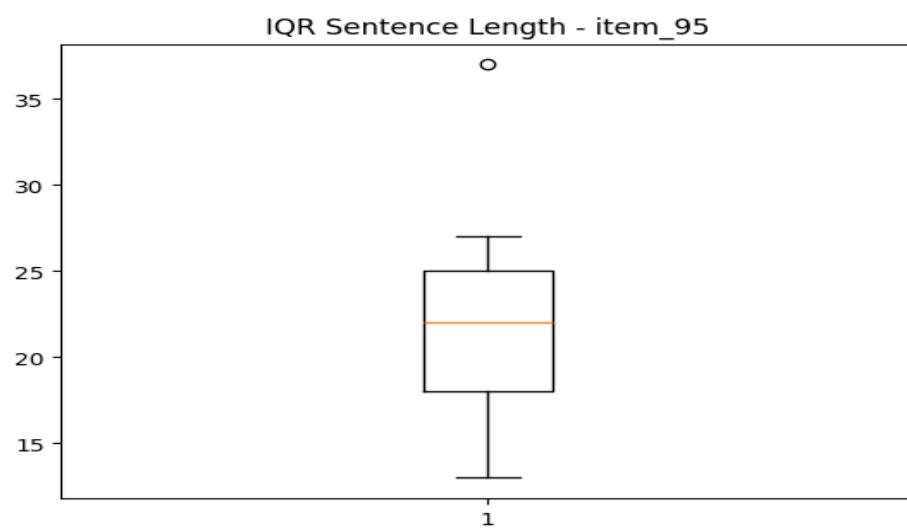
sentence_IQR_item_93.png



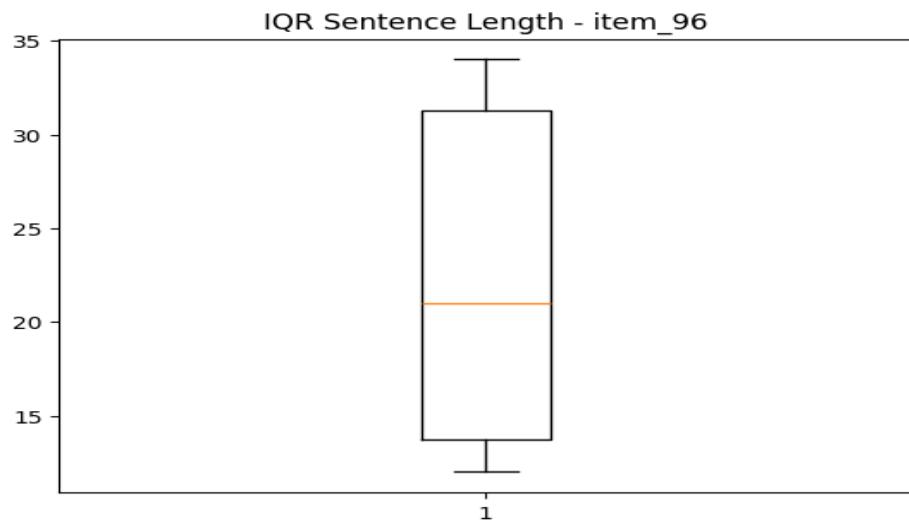
sentence_IQR_item_94.png



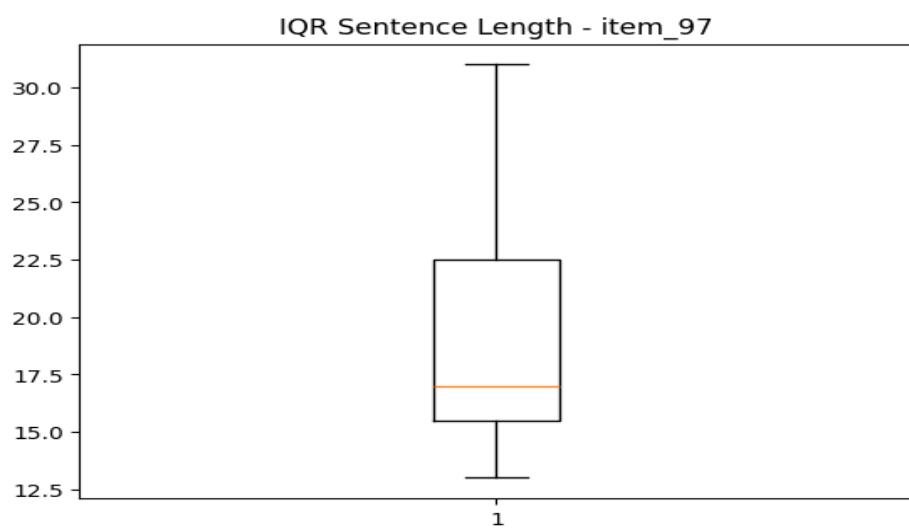
sentence_IQR_item_95.png



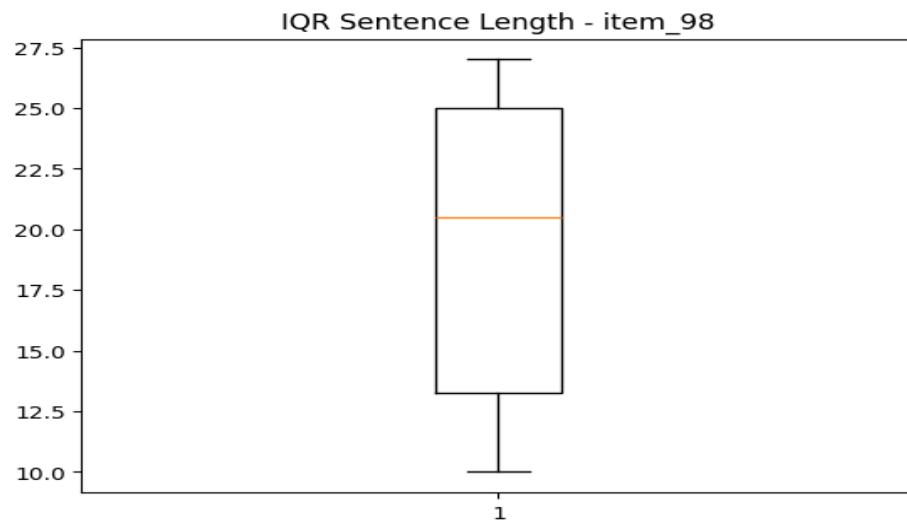
sentence_IQR_item_96.png



sentence_IQR_item_97.png



sentence_IQR_item_98.png



sentence_IQR_item_99.png

